Graduate Theses, Dissertations, and Problem Reports

2010

Exploring the utility of a GIScience approach to modeling invasive species: A case study of Ailanthus altissima

Arthur Elmes
West Virginia University

Recommended Citation

Exploring the utility of a GIScience approach to modeling invasive species: A case study of
Ailanthus altissima

Arthur Elmes

Thesis submitted to the
Eberly College of Arts and Sciences at
West Virginia University in partial
fulfillment of the requirements
for the degree of

Master of Arts
in
Geography

Approved by

Timothy Warner, Ph.D., Chair
James B McGraw, Ph.D.
Jamison Conley, Ph.D.
Rick Landenberger, Ph.D.

Department of Geology and Geography

Morgantown, West Virginia
2010

Abstract

Modeling the spread of *Ailanthus altissima* airborne seeds using remotely sensed data and agent-based modeling

Arthur Elmes

This thesis investigated the potential for integration of remotely sensed and GIS data into an agent-based modeling environment in order to model seed dispersal and subsequent establishment of windborne seeds. In order to explore the applicability of agent-based modeling to predicting seed dispersal, a case study was carried out using the representative example species *Ailanthus altissima*, an invasive tree found throughout North America's temperate regions. Seed movement was modeled in two stages, primary and secondary dispersal; primary dispersal was calibrated using existing field data, while secondary dispersal was calibrated only qualitatively. Establishment potential was accounted for probabilistically, based on landuse type. Environmental controls on seed movement and establishment were accounted for with several remotely sensed datasets. The general model characteristics and structure are representative of a potential class of predictive models that incorporate raster data and vector-based seed movement. Agent-based modeling provides a link between raster and vector data and processing methods, and is therefore a potential tool for projects involving both raster and vector data types as well as vector processing. Because seed dispersal and establishment modeling benefits from incorporating both of these data types, it was found that the agent-based approach provided an appropriate framework for modeling the phenomenon, while further research is necessary to fully parameterize and field-validate the model.

Acknowledgments

I would like to sincerely thank everyone who helped me complete this thesis. I especially wish to thank my advisor and committee chair, Dr. Timothy Warner, for his tireless work, insightful suggestions, and helpful guidance. Without his devotion, I would not have been able to complete such a daunting task as this. I would also like to thank the other members of my committee: Drs. Rick Landenberger, Jamison Conley, and James McGraw. Their ideas and critique helped to shape and guide this project, and they continually showed personal interest in my work. For providing data, I thank Drs. Landenberger and McGraw, as well as Alyssa Hanna and Matt Kaproth.

Finally, I would like to thank my friends and family, who lent me their support and ideas. I especially thank my father, Dr. Gregory Elmes, for his incisive critiques and ceaseless encouragement.

Table of Contents

List of Figures

List of Tables

1. Introduction

Invasive exotic species pose a serious ecological threat, as they have been shown to limit biodiversity and disrupt ecosystem function in both managed and natural environments (Wilcove *et al.,* 1998; Sala *et al.,* 2000; Lodge and Shrader-Frechette, 2002). This presents an economic as well as an environmental problem, due to the substantial costs associated with managing these invasive species (Mack *et al.,* 2000; Pimental *et al.,* 2000). Invasions are geographically widespread and biologically diverse, occurring in a broad range of environments by many species of plant, animal, and microbe (Vitousek *et al.*, 1996). By facilitating the short- and long-distance transportation of individual invasive plants or their seeds, human activity has substantially increased the spread of exotic species both locally and globally (Heywood, 1989; Lelong *et al.,* 2007; Pauchard and Alablack, 2003). Thus, there is an increasing need for accurate tools for prediction of the spread of invasive species.

In order to address this need, various modeling approaches utilizing geographic information systems, remotely sensed data, and ecologically-based methods have been used to map and model the spread of invasive vegetation species (Peterson 2003; Dark, 2004; Hurley *et al.*., 2004; Brown *et al*., 2008). These studies are typically conducted at broad scales of analysis, and rely on either spatial autocorrelation and autoregressive models of ecological and environmental variables or spatial diffusion models to predict areas vulnerable to invasion. Spatially explicit prediction of species dispersal is typically limited to the population-scale, and thus seed or other vector transportation is accounted for only in an implicit, generalized fashion. In addition, the raster/vector divide limits models to either raster or vector processing environments, further restricting the capacity for these models to account for both environmental variables, which are typically collected in raster format, and vector-based movement, which more realistically models dispersal (Higgins *et al.,* 1996).

Agent-based modeling, which can accommodate raster-format environment and ecology data as well as vector-based movement, allows the creation of a large scale, spatially explicit dispersal model. This study explored the potential to integrate raster and vector data with dynamic vector-based processing methods in order to model seed dispersal. Whereas an implementation-ready model would require extensive field verification and parameterization, this study focused primarily on the conceptual construction of the model . The object of this study was to explore the utility of the geographic data and methods to model the ecological phenomenon of seed dispersal in a generalized way, while setting up a framework to incorporate more precise future parameterization. In other words, the goal was to examine the ability of GISscience and agent-based modeling to capture the fundamental nature of the phenomenon within the computer model, and to yield accurate and reliable predictions. Thus, the contribution of this conceptual model is intended to be in the area of geographic information science (GISc), rather than ecology, *per se*. Although field verification and parameterization were outside the scope of this

study, the resulting model nevertheless has value as a hypothesis generation tool and, if further parameterized, as tool for land and resource managers.

2. Literature Review

This project draws on developments in remote sensing, geographic agent-based models (ABMs), and environmental modeling. In this section, literature on the following topics will be discussed: modeling invasive species with geographic information science (GIScience), ecology of *Ailanthus altissima*, remote sensing and geographic information systems (GIS) integration, and agent-based modeling.

2.1 Modeling invasive species with GIScience

Using GIScience to map and predict the spread of non-native invasive species has been explored by numerous previous studies of fauna and flora, including gypsy moth (*Lymantria dispar*) (Spears *et al.*, 1991; Hurley, *et al.*, 2004), Malabar plum (*Syzgium jambos*) (Brown *et al.*, 2008), rambler rose (*Rosa multiflora*) (Huebner, 2003), and tree of heaven (*Ailanthus altissima*, hereafter *A. altissima*) (Call and Nilson, 2003; Huebner, 2003; Knapp and Canham, 2000). Typically, GIS-based invasive species research utilizes either exploratory spatial data analysis (ESDA) of ecological variables (Dark, 2004), ecological niche modeling (Peterson, 2001; Brown *et al.,* 2008), a geographic diffusion method based on spatial dependence and diffusion rate equations (Hastings *et al.*, 2005), or some combination thereof. In these analyses, GIS serves as the analysis platform, providing a spatial framework within which to model diffusion, analyze patterns of spatial autocorrelation, or to overlay the spatial extents of limiting environmental variables.

Exploratory spatial data analysis of invasive species relies on measures of spatial dependence between groups of environmental and ecological variables, such as rainfall, landcover, and topography. Dark (2004) drew on GIS and exploratory spatial data analysis (ESDA) to find spatial patterns and dependence in invasive plants at a coarse spatial scale. This type of research, which is representative of geostatistical ecological modeling, focuses on finding multivariate spatial autocorrelation between invasive species and physical and man-made features, which in this case consisted of elevation, road density, and native plant species-richness. Dark's (2004) study used well-established statistical models such as ordinary least squares (OLS) and spatial autoregression (SAR) to characterize the types of environment associated with invasive species. The ESDA approach benefits from the strength of its underlying statistical framework, but is limited in the field of invasive species research because it is purely empirical, and does not include a mechanism for the spread or dispersal of objects over space.

Along a similar vein, Haltuch and Berkman (2000) used environmental and ecological GIS data layers to examine the invasion of exotic mussels in Lake Erie. Spatial regression techniques were used to predict the spatial extent of the invasive species at four time periods. These spatial data were then analyzed in order to estimate the rates and spatial patterns of expansion. This study, as with other similar studies, was very effective for coarse scale analysis.

Ecological niche modeling, also a statistical approach, can be applied to invasion modeling using two basic steps. First, the ecological constraints to the establishment and growth of particular species are defined by experiment or from the ecological literature. These constraints include such variables as climate, nutrient regime, elevation, and any other relevant ecological variable.. The range of the species' tolerance for each variable is determined, and the geographical extents of these constraints are mapped. The intersection of these ecologically defined extents represents the area potentially vulnerable to invasion by the species in question, irrespective of the current spatial distribution of the existing populations. Then, the current known extent of the species' distribution is mapped and juxtaposed with the potentially invasible area, allowing spatially and ecologically defined analysis of invasion (Peterson, 2003).

The assessment by Brown *et al*. (2008) of the threat posed by the invasive *Syzgium jambos*, illustrates the application of ecological niche models to invasive species research. First, Brown *et al.* (2005) used landscape-scale environmental variables, such as rainfall and leaf-area-index, to create an ecological niche model that predicted environmentally suitable habitats for *Syzgium jambos*. Second, they estimated rates of increase for each of seven populations of the plant, thus establishing the relative danger of spread from each location. By combining information concerning habitat suitability and the growth rates of already established populations, the study mapped the overall risk of invasion. Although clearly relevant for population-level management, this study dealt with relatively coarse data and simple assumptions. This approach inherently aggregates to broad geographic scales, and thus may be less helpful in predicting at the local scale the spread of invasive species from source individuals. Additionally, ecological niche modeling shares with ESDA the limitation that it has no mechanism for predicting diffusion through space.

In order to account for the process of diffusion, spatially explicit diffusion models have been proposed for various invasive species and in various geographic regions (Liebhold and Tobin, 2008; Liebhold *et al.*, 1992; Miller and Matlack, 2010). These models seek to predict the spread of invasive species based on historical ecological data along with modified diffusion formulae, Higgens *et al*. (1996) developed the Spatially Explicit, Individual-Based Simulation (SEIBS) model, which accounts for fire survival, adult fecundity, age of reproductive maturity, mean dispersal distance, and fire return interval. The model produced estimations of spread of an invasive pine species (*Pinus* spp.) that were evaluated to be reasonable in the context of empirical data (Higgens *et al.*, 1996).

Modeling the spread of gypsy moth (*Lymantria dispar*), Liebhold *et al*. (1992) used a spatial diffusion formula which accounted for diffusivity, intrinsic rate of increase, and organism density. The model, a derivation of Skellam's 1951 model, was parameterized using field data. The authors concluded that this model was inadequate in accounting for terrain heterogeneity, or the complexity of population growth, and suggest that more complex models, accounting for

anisotropic and/or non-Gaussian movement, discrete time steps, and stochasticity, would be more appropriate.

2.2 Remote sensing and GIS integration

To varying degrees, the above mentioned techniques for modeling the spread of invasive species rely on the tools and methods GIScience. The two major components of GIScience, GIS and remote sensing, provide a platform on which to collect and analyze geographic data, and a means of collecting and analyzing environmental data, respectively. However, these components are based on two conceptually different data structures: remotely sensed data is most often raster format, while most GIS programs work most effectively with vector data (Longley *et al.*, 2005). Although integration between remote sensing and GIS has increased markedly in recent decades, the fundamental differences between raster and vector data types have not been fully reconciled, and processing methods remain fairly disparate. For example, a vector-based GIS platform may allow for the simultaneous display or raster and vector datasets, but analysis is typically limited to one data type at a time, thus requiring a conversion process, which introduces uncertainty (Congalton, 1997). Remote sensing research tends to be dominated by pixel-based analyses, as the dominant data format is raster grid, whereas GIS research is typically based on an analysis of vector-defined features and their attributes. Recently, increased data and methods interoperability between the two fields have allowed more seamless raster/vector analysis and more unified raster/vector data structures, both of which have increased the utility of GIScience as a whole (Merchant and Narumalani, 2009).

The Netlogo agent-based modeling platform provides an example of a program that overcomes the raster/vector divide to some extent. The program incorporates the disparate data types as a part of its basic structure, as the mobile agents, which are vectors, and move through vector-defined geometric space, can interact with a gridded environment, which is a raster. Thus, no conversion between data types is necessary, and each data type operates under its native topological rules.

2.3 Agent-based modeling

Agent-based modeling is a way of predicting the movement of many individual, independent 'agents' over a plane or environment (Jennings, 2001). Agent-based models typically consist of two elements: mobile agents and a stationary background. The environment through which the agents move is constructed as a regular tessellation of square cells, and is essentially a virtual topography with potentially spatially differentiated values. These values can serve as movement 'costs' to the mobile agents, and can define agent-agent interaction rules. The relative ease with which agents can pass over the environment is determined partially by these movement costs, and partially by the agents' own attributes, which can be altered over model-time as differentiated environment is covered. It is important to note that the model's environment does

not inherently represent absolute geographic space, but rather can represent relational or even purely abstract space. For example, the environment across which agents move has been used to represent micro-scale space of biological systems (Bailey *et al.,* 2007; Johnson, 2009), relational organization space (Epstein, 2002), and abstract decision space (Joyce *et al.,* 2006; Earnest, 2008). Agent-based models of absolute geographic space are facilitated by the use of pre-defined geographic coordinate systems, which can be incorporated into the model's structure by defining the unit of angular measure, the standard meridian, and the datum. A model that incorporates these variables into the environment produces spatially explicit results.

Agents can be homogenous, that is, all have the same number and type of attributes, or they can have different attribute types or starting values. Theoretically, any number of agents with any number of attributes could be used, but hardware and software limitations tend to restrict the numbers to the hundreds or thousands of agents. The search for complex, informational patterns created by relatively simple but numerous agents is at the heart of agent-based modeling (Jennings 2001; Davidsson, 2002). It is important to note that a defining characteristic of agent-based modeling is that it is inherently decentralized, and there is no overarching control system or mechanism (Sycara, 1998). This bottom-up structure allows the model to create complex predictions using relatively simple input parameters, potentially displaying emergent patterns that could be missed by traditional, top-down models.

Agent-based modeling has been used in a number of research and practical applications, primarily within sociology, anthropology, and computer science, but also within hydrology and geography (Jennings 2001; Axtell *et al.*, 2002; Bonabeau, 2002; Epstein, 2002; Brown *et al.*, 2008; Reaney, 2008). Epstein *et al.*'s (2002) work on modeling civil violence and Axtell *et al.*'s (2002) reconstructions of ancient Anasazi population trends show how agent-based modeling has been used within social science to provide a concrete, quantified analysis framework, while Jennings (2001) and Woldridge (1997) address the application of agent-based modeling to the development of complex software systems.

Within social and computer science, agent-based modeling is often used to depict or model non-literal, more conceptual environment, with considerable emphasis placed on the importance of interactions between agents. For example, the civil violence model created by Epstein *et al.* (2002) depicted complex agents, representing pro- and anti-law-enforcement persons, interacting in an undifferentiated environment. The resulting cluster patterns were not geographic, in that they represented no actual area in reality, but simply showed how, all things being equal, groups of persons organize against civil authority.

Jennings (2000), describing the utility of agent-based modeling for the design of complex software systems, notes two major pitfalls of any such system: the outcome of this type of system is inherently unpredictable, and, a related problem, prediction of the behavior of the model as a whole based on the parameters and constituent entities is nearly impossible due to the

strong likelihood of emergent behavior. In addition to these issues, agent-based modeling of complex systems can be difficult due to computational limitations.

Related to agent-based modeling is the technique known as cellular automata, which abstracts the environment as a grid whose cell values change to represent the spread of spatial phenomena (Cole and Albrecht, 1999). Whereas agent-based modeling entails two types of entity, *i.e.* agents and background environment, cellular automata requires only grid cells, and is therefore somewhat simpler, both conceptually and computationally. However, agent-based approaches better account for complex systems, especially with multiple types of agents or environments. One advantage of agent-based over cellular automata is that the former incorporates mobile agents, which are not constricted to grid-based movement, but rather can move through euclidian space.

To a limited extent, both agent-based modeling and cellular automata have been used to predict the spread of invasive species, particularly for macro-scale analyses and/or to determine parameter significance. For example, in a sensitivity analysis, Cole and Albrecht (1999) investigated the contribution of various parameters to the spatial patterns of invasive species predicted by their cellular automata model. The authors discuss the impact of different parameterizations and different scales of analysis, noting the difficulties of selecting between stochastic and deterministic parameter values. In a more abstract, conceptual model, Suyamto (2010) modeled interspecies plant competition and dispersal across a variegated environment. This model did not predict the spread of any particular species, *per se*, but instead focused on modeling the effects of agent-to-agent, *i.e.* plant to plant competition, and the resulting pattern of dispersal. Like many agent-based models, this model employed a highly abstracted representation of the dispersal environment, again highlighting the focus on inter-agent interactions.

Parks *et al*. (2005) implemented the Argus Invasive Species Spread Model (AISSM), which is an agent-based model created within the AnyLogic modeling environment and capable of modeling several different invasive insect species at fairly coarse scales and large time-intervals. The model, which accounted for such variables as maximum distance of invasive species movement, minimum detectable population, and detection and eradication times, produced useful results on a statewide scale.

Applications of agent-based modeling within social or computer science tend to focus primarily on the attributes, behaviors, and interactions of the agents, rather than on the interaction of agents with the environment. Seed distribution modeling, on the other hand, relies much more on agent-to-environment interactions, and does not focus on interaction between agents. This makes the overall system considerably less complex than conventional uses of agent-based modeling, such as those mentioned above. This lack of complexity potentially allows a focus on accurate and precise calibration of the relatively limited number of model parameters.

In terms of modeling the spread of invasive species, the primary advantage of agent-based modeling is that it can incorporate both complex environmental data and a mechanism for spatial diffusion. As discussed above, geostatistical and ecological niche approaches to predicting the spread of invasive species incorporate complex models of the environment, but have no explicit accounting for spatial displacement. On the other hand, spatial diffusion modeling, which does explicitly account for spatial displacement, lacks the environmental complexity of the other two approaches. Agent-based modeling can potentially incorporate the advantages of each approach.

2.4 *Ailanthus altissima*

A. altissima was first introduced into the United States from Asia via Europe in 1784 as an ornamental plant (Merriam, 2003; Webster *et al.*, 2006), and was widespread in the state of Virginia by the 1880s (Hu, 1979; Call and Nilsen, 2003, Heubner, 2003). It displays extremely rapid growth and maturation rates, and can establish on compacted or otherwise marginal soil, making it a serious competitive threat to native shrub and tree species (Knapp and Canham, 2000; Webster *et al.*, 2006; Kowarik and Saumel, 2007). It is the seed production and dispersal of this species, however, that poses perhaps the greatest threat, as female trees produce up to 300,000 wind-carried seeds within three years of establishment, which are capable of establishment in a range of ecological conditions and tolerate stress well. Additionally, this species is capable of spreading rhizomatously (Hu, 1979; Call and Nilsen, 2003), forming extensive clones from root stumps. Thus so it is essential that either the entire root system be extirpated during eradication, or that sufficient herbicide is used to ensure complete eradication (Burch and Zedaker, 2003).

Although *A. altissima* is typically confined to open, high sunlight locations, including open fields, highway rights-of-way, and developed or built-up lands (Merriam, 2003; Webster *et al.*, 2006; Kowarik and Saumel, 2007), it is capable of establishment in forest margins and even, occasionally, forest gaps, provided that the gap size is large enough (Korwarik, 1995). Once a colony has become established, it tends to develop pure stands by out-competing the native understory vegetation (Burch and Zedaker, 2003). *A. altissima* may establish in tree-fall canopy gaps and similar forest locations, but typically remains competitive only in early successional stands, and is therefore much less of a threat in mature forest settings. Although it is occasionally found in mature and old growth canopy gaps (Knapp and Canham, 2000), it is primarily a threat in disturbed, low or zero canopy density areas (Landenberger *et al.*, 2007).

A. altissima seed dispersal occurs in two stages, called primary and secondary dispersal, respectively. Primary dispersal includes seed movement from the tree to the surrounding environment, and secondary dispersal includes all subsequent seed movement, including terrestrial and aqueous seed transport (Johnson and Fryer, 1992; Kaproth and McGraw, 2008). Seed dispersal is controlled by wind and environmental structure (Landenberger *et al.*, 2007).

3. Aims

A major constraint to modeling seed movement across a differentiated environment is the mismatch between the vector-based movement of the seeds and the raster format of the environmental data layers. This project sought to overcome this limitation by employing an agent-based modeling framework, which incorporates both data types and allows parameterization of seed movement in order to match field measurements. The primary aims of this project were as follows:

> i. Create an agent-based model of primary and secondary dispersal of wind-borne seeds, parameterized to qualitatively match field data. Full quantitative field verification and parameterization are beyond the scope of this project.

> ii. Critically assess the effectiveness and accuracy of the agent-based model, the usefulness of its individual data layers, and the overall model architecture.

>> a) Run the model and qualitatively calibrate it to the *A. altissima* seed dispersal data created by Landenberger *et al*. (2007). An acceptable dispersal pattern, i.e. a positive result, is judged on a qualitative assessment of plausibility based on field evidence and expert testimony.

>> b) Address model strengths and weaknesses in terms of the synthesis of data types and processing methods in comparison to non-ABM invasive species models.

> iii. Evaluate the utility of ABM as a platform for raster/vector interaction for geographic phenomena.

> iv. Investigate the potential of this modeling environment for management applications, hypothesis generation, and future research on invasive plants and seed transport.

The focus of this study was to illustrate the capability and flexibility of an agent-based approach in modeling the biological phenomenon of seed dispersal, rather than to create an implementation ready, field-verified model of *A. altissima*, *per se*. This species was used as a case study, as its seeds can be seen as representative windborne agents, and because the threat posed by invasive species illustrates well the need for accurate and precise models of seed dispersal and establishment.

4. Study Area

The study area, as shown in figure 1, consisted of an approximately 17 square kilometer area around Morgantown, West Virginia (approximately. 39° 37' N, 79° 56' W). The Morgantown area was selected primarily due to the availability of environmental datasets and field-collected *A. altissima* location data. In addition, the Morgantown area is representative of a broad range of topographical, vegetation, and human-altered micro-environments that can be found in the greater Appalachian region, including dense forest, rocky outcrops, grassland, urban and suburban areas, and highway infrastructure.


Figure 1: The study site. This figure shows the field collected *A. altissima* tree locations atop the NAIP aerial imagery. The imagery was collected in August of 2007 and comprises green, red, and near-ifrared bands, which were colored blue, green, and red, respectively, for display. Note that this image shows the same extent as figures 3 through 5.

5. Methods

The process of adapting and simplifying the real-world biological system of seed dispersal to an agent-based computer model has two stages: selecting the salient controls, *i.e.* the environmental layers, and parameterization of the environment and agents. Below is a discussion of the datasets chosen to represent the environment through which the modeled seeds moved, followed by an explanation of the structure of the agent-based model itself.

5.1 Data layers

The differentiated landscape across which the modeled seeds moved and germinated was constructed from four data layers, each representing a different control on seed movement and/or seed establishment potential on a per-pixel basis. These layers were chosen because of both their availability and because of their effect on seed transportation. In addition to the four environmental layers, a fifth layer representing the locations of *A. altissima* trees was used. All datasets used in the model are listed in table 1.

Table 1: Data layers used in model.

Data Layer	Data Collection Date	Source	Resolution	Contribution to Model
Landsat Imagery	2005	NASA/USGS	30 Meter	Landuse type
NAIP Imagery	2007	USDA/FSA	1 Meter	Landcover type
Lidar Derived Surface Roughness	2003	Spectrum Mapping, LLC	Approx 9 pulse returns per meter	Surface roughness
Digital Elevation Model	2003	USGS/WVSAMB	3 Meter	Topography
A. altissima **locations shapefile**	2004	Landenberger *et al.*, 2009	Field collected	Tree locations

A short description of each layer used in the model follow.

5.1.1 Landuse

Landuse classes were derived from a maximum likelihood classification of 30 meter resolution Landsat 5 satellite imagery, path 17, row 33, acquired in September 2005. It was assumed that each generalized landuse class was associated with a slightly different average influence on the likelihood of any given seed to germinate and establish. Therefore, each class defined the value of the 'establishment probability' attribute for the agent-based model's environment. The moderate resolution (30 m) of this dataset allowed a generalized interpretation of landuse, in contrast to the fine resolution of the aerial imagery used for land*cover* classification. The larger Landsat TM pixels were assumed to be better suited to capture landuse at the scale of analysis

used by this model, whereas a finer-scale was assumed to be more useful for representing land cover. Table 2 lists the landuse classes used:

Table 2: Landuse classes and associated establishment parameter values (higher values indicate greater likelihood of establishment).

Landuse Class	Establishment Value in Model
Open grasslands/field	5
High density development, including urban areas	4
Low density development, including residential areas	3
Forest	2
Water	0

Although a range of ecological and environmental factors affect *A. altissima* germination and establishment potential (Kota *et al.,* 2007; Landenberger *et al.,* 2009), human practices also strongly contribute to the likelihood of establishment. Because the landuse classification data layer can be used to infer human interference or facilitation of establishment, it was used as the control on establishment likelihood. This layer represents only a coarse-scale control on seed establishment; additional data layers of this type would be needed to fully parameterize establishment likelihood. The distinction between germination and establishment was not explicitly made; further prediction of tree success was not considered, and thus germination and establishment were combined into one parameter, henceforth referred to as 'establishment'. The establishment values in table 2 were arbitrarily chosen based on the landuse inferences discussed below.

The low density development class, which comprised primarily residential or low density commercial areas, contained a mixture of concrete and vegetation pixels. This landuse type was inferred to represent some degree of artificial vegetative maintenance, *i.e.* selective weeding. Although *A. altissima* is not immediately and universally removed from every residential or commercial property, it was assumed that these areas undergo more selective weeding than heavily developed areas or barren/undeveloped lands. Therefore, the probability of establishment low density development areas was assumed to be slightly lower, even though they present a suitable ecological environment for establishment. *A. altissima* has been shown to have establishment success in urban and suburban (*i.e.* 'high density development') areas (Hu, 1979,; Webster *et al.*, 2006), due in part to the species' ability to utilize resource-poor soils and other unfavorable growing conditions. Thus, this class was given a higher probability for establishment than the low density class. *A. altissima* has not been shown to successfully invade mature forests (Knapp and Canham, 2000); thus these areas were given almost zero probability of establishment. Pixels with a high probability of establishment were given values near five, while pixels with a low probability were given values near one. Establishment probability values were assigned to these classes conjecturally, and are not supported by field data. This data layer

was included in the model in order to highlight the capacity of the model to predict establishment, provided sufficient field data and parameterization.

It should be noted that the Landsat classification did not generate many forest pixels relative to the other classes, tending to bias towards the low density development class. This can be attributed to the complex spectral composition of the low density development class, as well as the fact that many 'low density development' training areas included mature trees.

5.1.2 Landcover

This data layer distinguished classes based on vegetation and impervious surfaces using a supervised classification of National Agriculture Inventory Program (NAIP) 1-meter color infrared imagery, which was acquired in August of 2007 with Z/I DMC metric digital mapping cameras operated by Photo Science, Inc. This high-resolution imagery allows complex spatial patterns of surface type to be measured at a scale appropriate to the model's needs. The classification technique used was the same maximum likelihood supervised classification tool as that used for the landuse classification. Table 3 lists the landcover classes used.

Table 3: Landcover Classes. Note the arbitrary high value for the water class, which ensures that water presents an absolute barrier to seed movement within this model.

Landcover Class	Friction Value in Model
Concrete, asphalt, and other impervious surfaces	1
Grass	3
Forest	10
Water	100

Impervious surfaces were assumed to provide the least resistance to seed movement, and were thus modeled with low resistance to seed movement. The two remaining vegetation classes, grass and forest, were given higher resistances to seed movement. Field data collected in 2006 by McGraw and Kaproth showed that during secondary dispersal in a forest setting, the vast majority of *A. altissima* seeds move one meter or less (Personal Communication). Thus, forested areas were modeled to allow only limited seed movement. Although Kaproth and McGraw (2008) and Kowarik and Saumel (2008) showed that waterborne secondary dispersal can be an important aspect of seed transport of *A. altissima*, this study assumed that any seeds that moved onto a water pixel were incapable of further movement. Although it is likely that these seeds would move to a distant location, the limited extent of the study area precluded prediction of stream-based dispersal of seeds. Thus, the water class was given an arbitrary high friction value in order to immediately stop any seeds which landed on a water cell.

5.1.3 Surface roughness

The high-resolution imagery described above provided an estimate of resistance to seed movement based on landcover type. To supplement this, lidar data was used to measure surface roughness directly. Lidar data of the study area was collected on October 2003 by Spectrum Mapping, LLC. The data were collected at a laser pulse rate of 35 KHz, a scan rate of 25 Hz, and at an altitude of 2,000 m, yielding approximately one return per square meter, and thus providing adequate resolution to differentiate between shrubs, trees, and 'bare-earth' features, such as grass, rocks, and other dense, low-lying objects.

The lidar data was interpreted by the Fusion software platform, developed by the United States Forest Service specifically to analyze vegetation structure (http://www.fs.fed.us/eng/rsac/fusion/). Each laser pulse potentially is associated with multiple returns to the sensor, and by comparing the number and order of these returns, the program distinguishes tree canopies and other objects from the ground. Multiple-return pulses typically indicate some form of vegetation or above-ground obstacle, as the beam is reflected off various branches, leaves, or other obstacles as it penetrates the canopy. The various returns are stored discretely within the dataset, meaning that the analysis software can separate first-and-only and last-of-many returns from all others, which enables the creation of a bare-earth model that includes only the surface features relevant to seed movement. The bare-earth or ground surface model (GSM) was created from the bare ground subset of the original point data in order to model surface roughness in raster format, so that it would be compatible with the other environmental data layers in the model. The goal was to model roughness at a fine scale, so the GSM created was smoothed as little as possible during creation. In other words, as many points from the lidar ground subset as possible were included in the GSM raster.

Once the GSM was created, roughness was calculated using a vector ruggedness measure, which measured raster ruggedness by calculating variation in relief and aspect within a neighborhood of elevation grid cells, accounting for both slope and aspect in one measure (Sappington *et al.*, 2007). The end result was a raster data layer that accounted for terrain roughness at a scale commensurate with such obstacles to potential seed movement as tall grass. Primarily, however, this layer successfully highlighted borders between surface roughness *type*, such as the boundary between a field and a road. Artificial features, such as buildings and walls, were starkly delineated by this dataset, which had the effect of limiting seed movement over such obstacles. It should be noted that this layer showed a positive correlation between slope and roughness estimation, as natural elevation changes such as steep hills or gullies were given higher roughness values, irrespective of their micro-scale roughness. Because topography also was accounted for with another data layer, this presents some redundancy, which could potentially bias topography above other seed movement controls. To circumvent this problem, the roughness layer was given less weight in determining the control to seed movement.

5.1.4 Topography

A digital elevation model DEM was used to account for the effect of slope and aspect on seed transportation during both primary and secondary dispersal. The photogrammetrically produced three-meter DEM used was created by the West Virginia Statewide Addressing and Mapping Board (WVSAMB) in 2003, and conforms to the American Society for Photogrammetry and Remote Sensing (ASPRS) standards for both vertical and horizontal accuracy of three meters.

5.1.5 *Ailanthus altissima* tree locations

The locations of individual female *A. altissima* trees were obtained in ESRI shapefile format created by Landenberger *et al*. (2009). This file contains 107 total records in the two semi-urban study plots which provide a range of slope, aspect, and landcover types. These points were collected in the field in January and February of 2004.

5.2 Layer preparation

In order to ensure consistent extents and grid cell sizes between analysis layers, each of the above raster datasets was subjected to several pre-processing steps before input into the ABM. First, each data layer was projected into the Lambert conformal conic projection, which uses the North American Datum of 1983, has standard parallels at 20 and 60 degrees north, and whose linear unit is the meter. This projection was chosen purely because it is one of the few supported within the Netlogo environment.

Next, the landcover and friction rasters were combined into one composite layer using simple raster algebra. The combination of these two layers composed the 'friction' control on seed movement, and were thus taken as one entity in the model. The layers were combined on a pixel-by-pixel basis using a generic resistance calculation:

$$1/((1/\text{landcover}) + (1/(\text{roughness})))$$

After this raster calculation, a separate calculation was carried out to add the pixels classified as 'water' in the landcover image to the final friction layer. These pixels were given very high values, ensuring that water represented an impassible feature for the analysis.

Note that were field experiments carried out, the relative contributions of these layers, as well as any others capable of being represented in raster form, could be modified as much as necessary to best reflect reality. The model presented in this study makes the assumption that landcover type is more important than surface roughness, given the correlation between roughness and slope in the roughness raster.

Next, an arbitrary rectangular polygon was drawn around the study area, the edges of which were at least 800 meters from any given *A. altissima* location. Each raster dataset was clipped to this polygon, yielding rasters of identical size. The friction dataset was then resampled to two meter pixels. Smaller pixels were preferable for this analysis due to the scale at which the phenomenon to be modeled occurs, and two meters was the finest resolution that was both supported by the data and by the Netlogo platform, given file size constraints. Neither the landuse classification nor the digital elevation data were resampled, but retained their original resolutions of 30 and three meter pixels, respectively. Down-sampling of these rasters to match the two meters of the friction dataset would not alter or improve the model's function, and would only increase the computational demands of the model. Finally, the datasets were converted to ASCII grid format, as this is the only raster format that Netlogo accepts.

## 5.3 Agent-based model structure and parameterization

At the most basic level, an agent-based model consists of a set of independent, autonomous agents, which can interact with each other and/or with the model's background environment, and whose rules for movement and interaction are pre-defined. The background need not be differentiated; in fact, many models focus entirely on agent-to-agent interaction, and thus have no need for a raster background (Epstein, 2002). These agent-agent models have been used to simulate a wide range of phenomenon, often far outside the purview of geography. In the model created for this study, however, the agents interact *only* with the background, which comprises multiple data layers, as described above. By establishing the basic rules of individual seed movement, the more complex patterns of overall seed dispersal were investigated.

Netlogo, an agent-based programming platform developed by Wilensky (1999), was used for all modeling carried out by this study. Although many other agent-based platforms are available, this freeware program was used due to its capacity to directly import GIS data formats. As with most agent-based modeling platforms, Netlogo places more emphasis on agent-to-agent interaction than on agent-to-environment interaction, but it can incorporate multiple environmental layers simultaneously. Netlogo allows both agents and raster grids, or patches, to be manipulated and defined, and incorporates both vector and raster data types in their native formats. This means that raster datasets are treated as grids of values, but agents are not limited to rasterized movement, such as rook- or queen-based movement. Rather, agents move in angular space, meaning that they interact with whichever pixel they are atop, irrespective of their position on that pixel. Additionally, Netlogo's GIS extension allows several key data interoperability steps, including shapefile and ASCII grid import and export, as well as the ability to incorporate the attributes of multiple raster datasets into the model. This tool also maintains geographic coordinate system and projection information, so that exported data can be examined and re-analyzed on other GIS platforms.

5.3.1 Agents

Within the model's environment, which is defined by the maximum extent of the raster datasets, the agents are generated at coordinates defined by the shapefile containing their field-measured locations. The number of agents generated at each location is specified beforehand, and can range from several dozen to several hundred thousand. Provided sufficient computer processing time, Netlogo is capable of modeling the hundreds of thousands of seeds produced by mature *A. altissima* individuals; however, much of the analysis was conducted with only thousands of agents at each location.

When the program is run, the agents first move radially outward from their starting locations and then stochastically, as controlled by the environmental data layers. This mimics the two distinct stages of seed dispersal: primary and secondary dispersal (Chambers and MacMahon, 1994). The former describes seed transportation from the tree to the ground, and the latter describes transportation from that point forward. Primary dispersal was modeled to match existing seed density-distance curves, but also took into account wind bias and topography. The model allows for both average wind direction bias and discrete wind-based dispersal events, such as storms. Average wind direction is modeled by dispersing the seeds radially, but with a normally random distribution around a direction, defined by degree. The standard deviation of this distribution allows the 'tightness' of the wind-bias to be defined. Discrete wind-based dispersal events are accounted for using the same method, but by creating different groups of agents for each starting location, each with a different dispersal pattern, defined by unique direction/standard deviation combinations. Hereafter, these discrete wind-dispersal events will be referred to as waves of dispersal. Elevation controls primary dispersal only to the extent that agents are biased against dispersing to locations above the assumed height, including elevation, of the tree from which they originated. In other words, if the elevation to which a seed would otherwise move is calculated to be higher than the seed's current location plus the height of the tree, the seed moves a shorter distance in that direction. The turbulance effects of topographic features were not modeled.
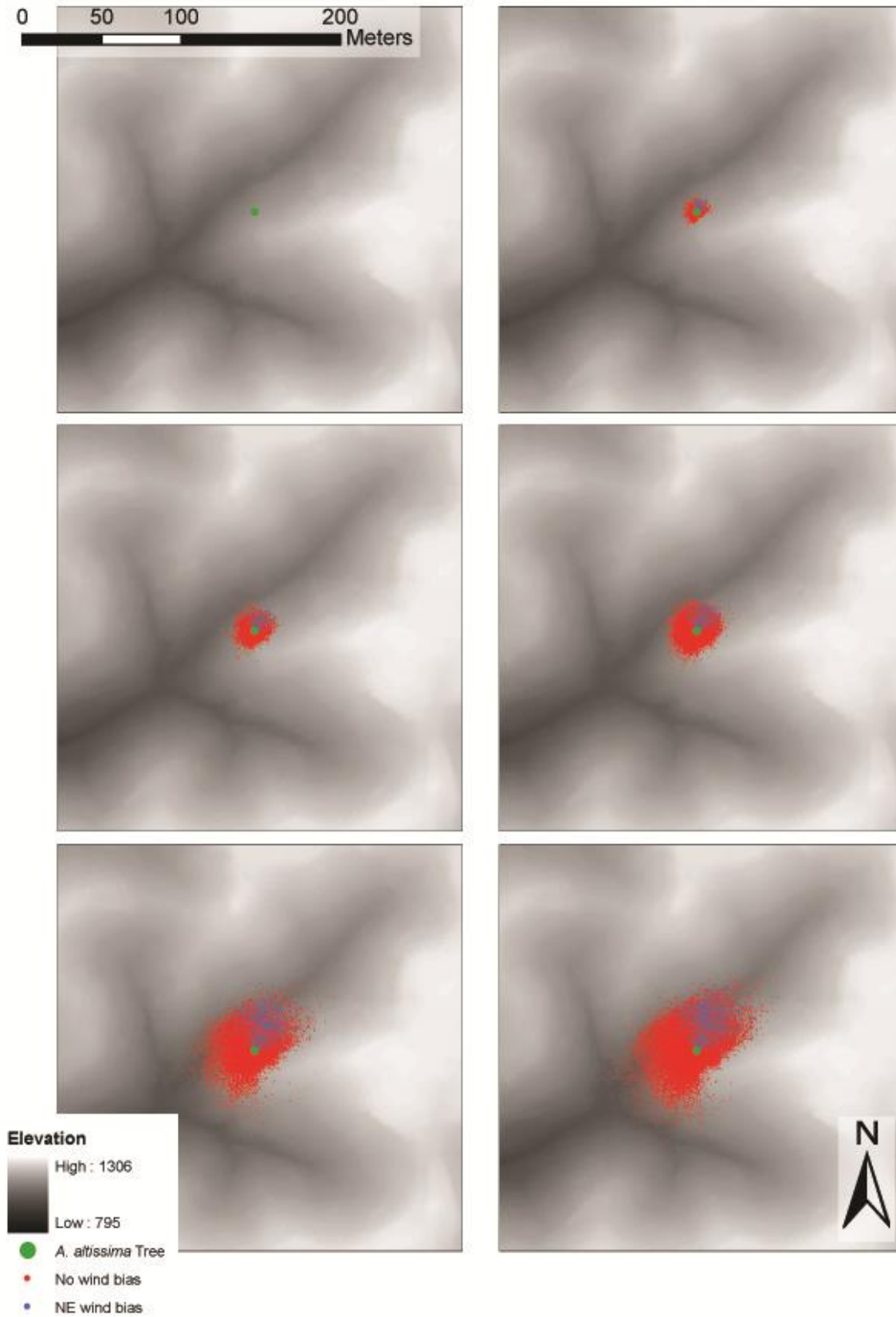
Figure 2: A series of snapshots of primary dispersal.

The following steps describe the modeled primary dispersal, once the agents were generated at

their starting locations. Note that the following steps apply to all waves of dispersal, but each wave disperses independently and serially.

Figure 2 shows a small subset of the total extent of the model. . This scene contains one *A. altissima* individual, and was thus selected purely for illustrative purposes. Near the center of the image is the tree from which all the agents originate. Agent color indicates the wave of dispersal to which that agent belongs: red agents represent a non-biased dispersal, while blue agents represent seeds subject to a discrete strong wind event with winds from the south-southeast. Such wind events are programmed into the model *a priori*, arbitrarily in this case. When the model is run, agents at the starting location were assigned a random heading using the random distribution assigned to their dispersal wave. For example, the non-biased (red) agents were assigned a random heading between one and 360, while a wind-biased group is assigned a random heading from within a random distribution of degrees with a mean of the predominant dispersal direction. Thus, a discrete wind event with winds to the northeast would be modeled as a random distribution of headings with a mean of 45 degrees. The standard deviation may also be specified to determine the 'tightness' of the dispersal; strong wind events have lower standard deviation, which focuses the distribution of headings and thus the dispersal pattern of seeds.

Once agent headings have been assigned, primary dispersal begins, as illustrated in panes (a) through (f) of Figure 2. Primary dispersal was modeled to match a set of field collected seed distance-density distributions (Landenberger *et al.*, 2007). Because the shape of the distance-density curves were somewhat site-dependant, the data from this study were averaged to create a composite density-distance function, which comprised data from four sites: two field sites, one forest-adjacent site, and one site adjacent to a recently harvested forest. Landenberger *et al*. measured seed density using seed traps along 11 transects per site, spaced at 10 meter intervals starting at the tree. In order to match this data, an exponential function was fit to the composite distance-density distribution, and was then used to predict seed densities at five meter intervals from zero to 200 meters from the tree. The proportion of seeds found at each distance was calculated, relative to the total number of sampled seeds. This proportion was used to determine how many of the modeled seeds, *i.e.* agents, would travel each distance, from zero to 200 meters. A random number generator was used to distribute seeds across each five meter interval, thus alleviating the otherwise stark concentric ring pattern. For example, the composite curve predicted that approximately 12% of all sampled seeds traveled five meters, 10% traveled 10 meters, 8% traveled 15 meters, *et cetera*. It was assumed that these proportions included seeds traveling anywhere within the five meter buffers of each area, thus 12% fell between zero and five meters, 10% within five and 10 meters, and so on.

Once the model had been run, the predicted dispersal densities were compared to the composite distance-density data derived from Landenberger *et al.* by measuring seed density within 1 meter square polygons set at the same intervals as Landenberger *et al.*'s transects. This was carried out
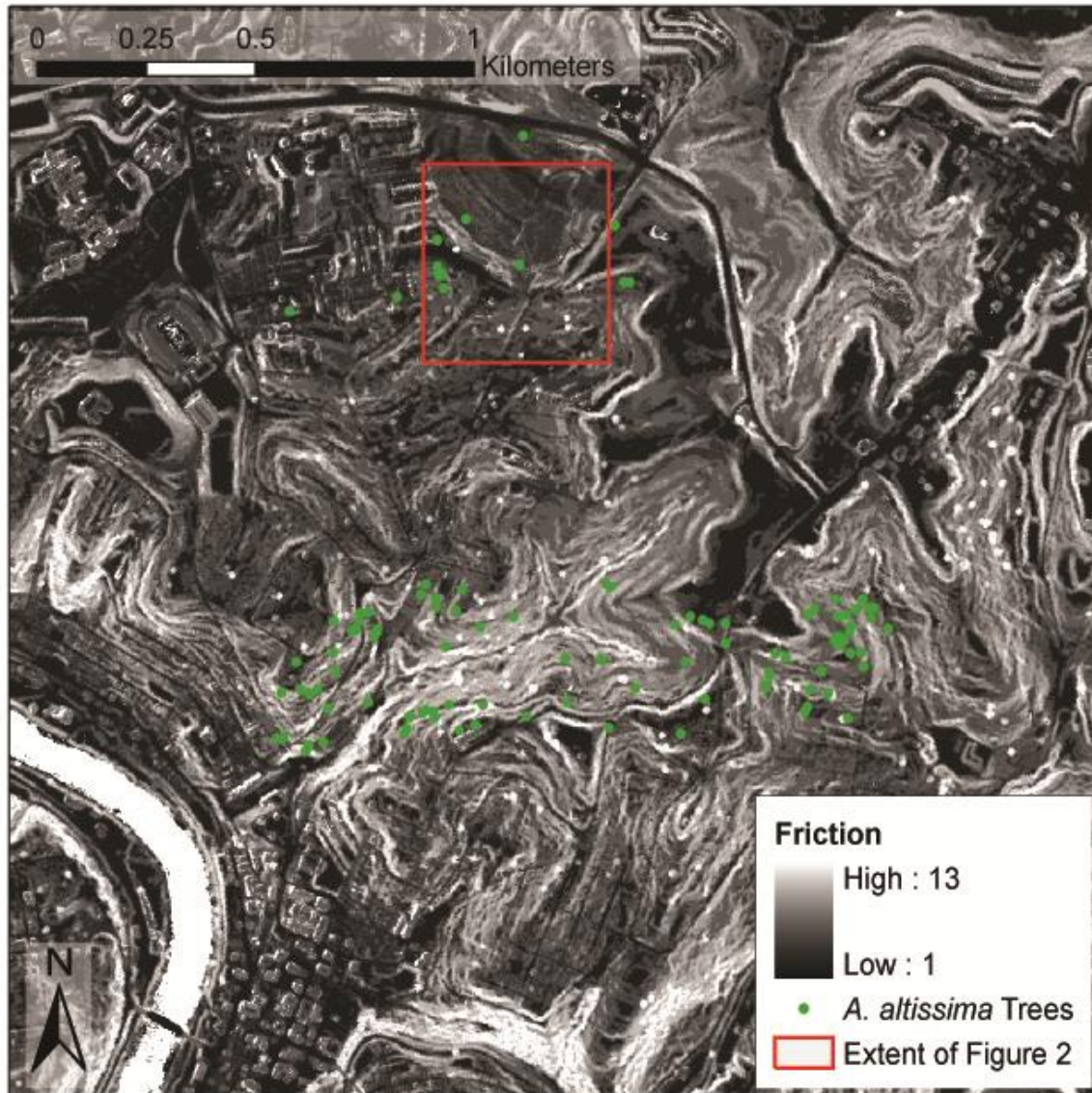
Figure 3: The entire study area, showing the friction surface and tree locations.

in a separate GIS, and was intended to mirror the field experiments as closely as possible.

Panels (b) and (c) of figure 2 show snapshots of primary dispersal in progress, and the fourth panel shows the final result. Note that in this illustrative example, the number of agents was reduced by an order of magnitude in order to more clearly show individual agents. Panels (e) and (f) show hypothetical strong-wind dispersal events, in which the blue agents are affected by these events. In panel (e), the blue agents represent a diffuse event with winds from the

southwest, while panel (f) shows agents affected by a somewhat stronger event with winds from the same direction.

In addition to *a priori* seed dispersal distance-density, seeds undergoing primary dispersal were assumed to be controlled by topography, and thus the elevation raster was invoked during this stage of the model. Before moving, each agent 'looked' ahead the number of units described above, and the elevation of this grid cell was compared to that of the seed's current location, plus an assumed 15 meters for tree height. If the elevation to which the agent is 'trying' to move was higher than the current elevation, the seed moved a shorter distance. This topographic 'check' was carried out at each step of an agent's movement, in order to ensure a consistent topographic control on seed dispersal. Figures 3 and 4 show the beginning- and end-points of primary dispersal in the context of surface friction.

For primary dispersal, agents were differentiated into classes by dispersal wave in order to facilitate wind-bias modeling and parameterization. Once primary dispersal was complete, the agents were undifferentiated from one another in terms of movement, and were thus all displayed in green during secondary dispersal modeling. Secondary dispersal was modeled as movement across the differentiated background consisting of friction and elevation controls. In order to ensure that every agent eventually ceased to move, the variable 'energy' was given to each agent after primary dispersal. It should be noted that this variable does not characterize any physical energy; it is simply designed to allow the agents to move terrestrially for a limited amount of time following primary dispersal. The starting value of this variable was equivalent for every agent, but it was decreased at each movement step by the friction value of its current raster cell. Additionally, agent movement could be stopped by establishment, the calculation for which is explained below. Agent movement ceased when the starting energy had been depleted by cumulative friction, or when they reached a pixel with suitable establishment potential.

Field data collected by McGraw and Kaproth indicate that during secondary dispersal in a forest setting, most seeds travel fewer than 10 meters (Personal communication). Of the sampled seeds, in fact, the average observed movement distance was approximately 0.2 meters in the summer and 0.01 meters in the winter. Because these distances are below the scale of this model's analysis, secondary dispersal within forested areas was restricted to steep slopes, which were observed to facilitate seed movement up to approximately 10 meters.

The model progressed in time steps, during each of which every active agent would move independently and serially. Model time steps do not correspond to actual time; they are purely a construct of the model. For each potential movement, an agent rotated a random number of degrees between one and 360, and then had the opportunity to move along that heading. The distance moved by an agent at a given time step was contingent upon the properties of the grid cell on top of which it was located. The composition of these grid cells is described in the following section.
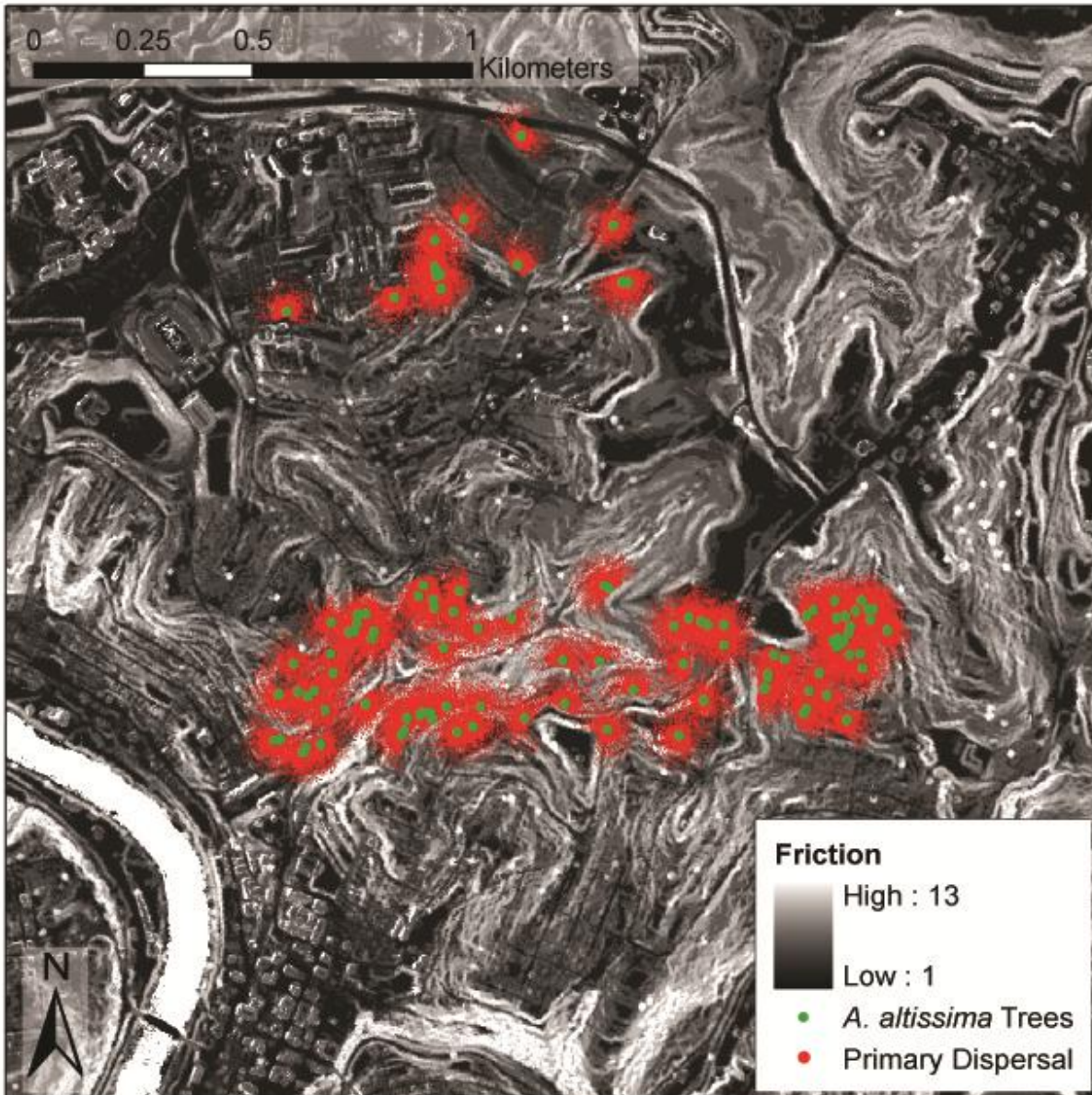
Figure 4: Primary dispersal. Note the locations of the *A. altissima* trees, marked in green.

5.3.2 Environment

The agents all represent the same physical phenomenon, *i.e.* dispersing seeds, and each agent is fundamentally identical, capable of the same domain of movements, and thus comparatively simple to model, once the basic interaction rules have been established. The raster datasets, on the other hand, represent a more complex aspect of this model, as they are designed to represent the real-world environment as it pertains to seed movement. The topographic, friction, and establishment controls modeled by the raster datasets were all accessed separately by the agents, in order to maintain the separate parameterization control in the model. Topographic control on seed secondary dispersal was modeled in precisely the same way as during primary dispersal:

any move requiring a positive elevation change beyond a small limit was ruled out, and the agent
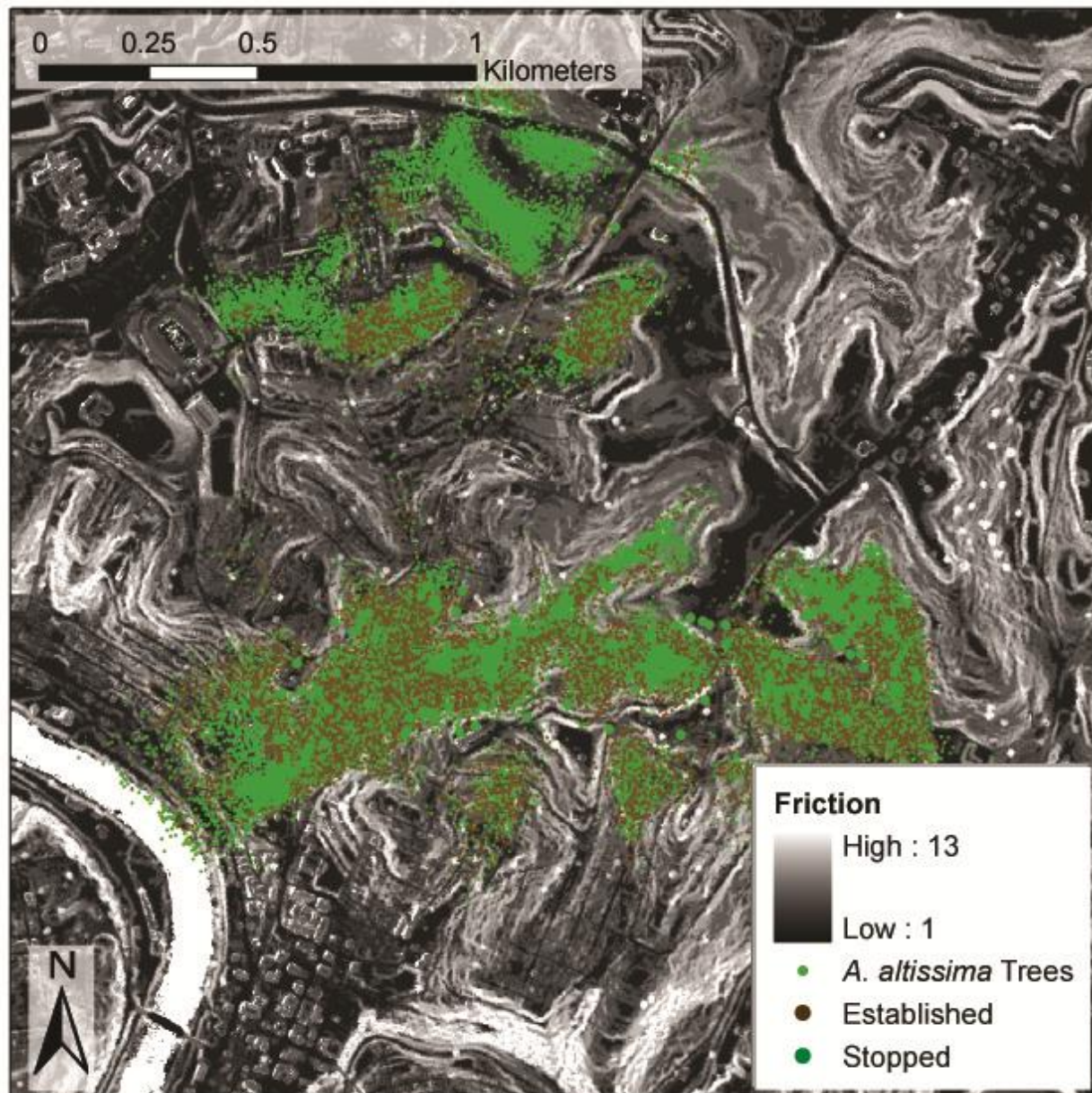
Figure 5: Secondary dispersal. Brown indicates that the agent has potentially germinated, while green indicates that it has simply stopped moving.

did not progress in that direction. The establishment potential layer, *i.e.* the landuse classification, was queried by the agent before each potential move in order to determine whether the agent would more likely germinate at the current cell or move to a new cell. This calculation was probabilistic, and was carried out at each movement step. The values in the establishment dataset ranged from one through five, with five being the most suitable for establishment. Because the larger values were more likely to be greater than the randomly generated number, the grid cells with these larger values were more likely to cause an agent to stop moving and

become established. Figure 5 shows the end result of secondary dispersal, including the germination calculation. The heterogeneous pattern of germinated seeds reflects the probabalistic nature of this calculation.

The combination of surface roughness and landcover type constitute the 'friction' layer, which constrains agent movement analogously to the conventional definition of friction, *i.e.* resistance to movement between two surfaces. The full extent of this layer can be seen in figure 3. When secondary dispersal begins, each agent queries the friction variable of its current grid cell before every potential move, and subtracted that value from its own energy value. When an agent's energy was equal to or below zero, that agent ceased to move. In addition to determining the number of moves allowed to an agent, the friction layer also determined how far the agent moved during each turn. High friction values yielded shorter movement distances, and *vice versa*. Therefore, agents which landed on high friction cells moved less far and stopped moving at an earlier time step than those that landed on low friction cells.

5.4 Summary of model structure

In summary, this model was constructed from mobile agents, which represented seeds, and a static background, which represented the environmental controls on seed movement. The mobile agents dispersed from the trees in two steps, primary and secondary dispersal, and were subject to a germination likelihood calculation during secondary dispersal. The background comprised surface friction, elevation, and germination likelihood rasters, and acted as a control on agent movement.

5.5 Scenario simulation

One of the key advantages the ABM is that it can be used as a hypothesis generation tool. In order to help elucidate the adaptability of this model under different distribution environment conditions or different initial starting location distributions, several artificial scenarios were created. The following simulations were carried out by manual editing of the environmental raster datasets or the *A. altissima* location shapefile, and thus represent only conjectural scenarios.

Figure 6: Dispersal of one tree in a clustered group versus that of the entire group.

The first simulation investigates the consequences of failing to exterminate one tree of a clustered group, by comparing the dispersal of the entire stand with that of a single tree from within that stand (Figure 6). A subset of the *A. altissima* dataset including only a clustered group of trees was clipped out of the original file, and the model was run for only those trees, producing a local dispersal prediction. Then, all of these trees were eliminated, save one, and the model was run again. Although the distribution of seeds from a sole tree is naturally much less widespread than from the entire group, the single tree was predicted to distribute seeds to a large portion of that area predicted by the original stand. Note that figure 6 shows the dispersal pattern on the background of the elevation control layer, which shows the effect of topography on dispersal.

In another simulation, the effects on seed dispersal by a hypothetical paved area were examined. This simulation could perhaps represent a large commercial development and parking lot, but the specific landuse change is less relevant than the landcover change: from grass and forest in the original, unmodified friction layer, to purely paved pixels in the modified version. Because paved surfaces represent comparatively low resistance to seed movement, new construction of this type represents a potential expedient to *A. altissima* invasion. This conjectural scenario illustrates the applicability of ABM for modeling the effects of potential changes to the physical environment. In order to create this simulation, a polygon representing the area of the hypothetical paved surface was digitized, rasterized, and added to the friction layer.
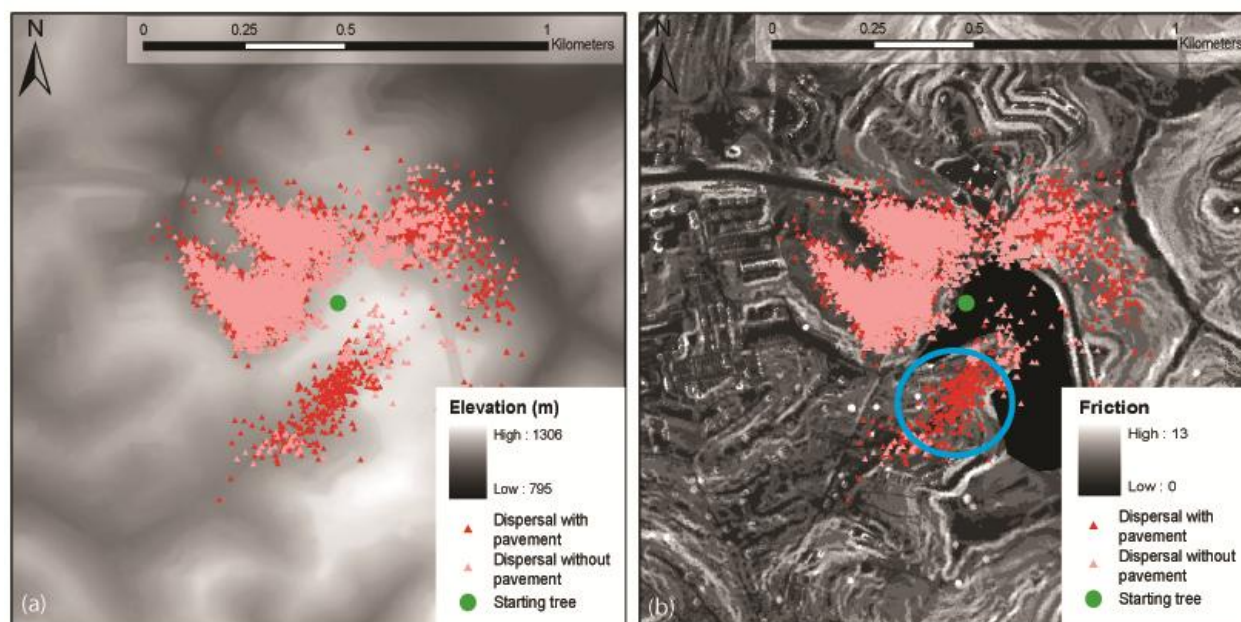
Figure 7: Dispersal from one tree under two different friction scenarios.

Figure 7 shows the final result of secondary dispersal from the same tree under both friction scenarios: the red agents have moved across the unmodified friction layer, while the pink agents have moved across the hypothetical new-pavement layer. Both images show the same dispersal event; panel (a) shows a hillshade of the elevation layer, while panel (b) shows the friction layer.

The black region to the southeast of the green point represents the hypothetical new paved area. The green point at the center of the image represents the starting tree, which is located on relatively high ground, as the hillshade shows. The right pane shows the same scenario, but on the background of the hypothetical friction layer. Because the paved surface represents little constraint on agent movement, very few of the agents that moved across the modified surface (colored red) stopped on the hypothetical paved area, whereas the agents that moved across the unmodified surface (colored pink) can be seen to cover this area sparsely, indicating that they were stopped by the friction of the unmodified surface. When the surface was modified to have a uniform, low, friction, many fewer seeds ceased to move on top of this area, reflecting the comparitive ease of movement over pavement versus vegetation. Importantly, the paved area allowed many more seeds to move to the area directly south of the origin tree (ringed in blue), which was almost empty of seeds in the no-pavement scenario One benefit of the agent-based approach is that it can highlight unintuitive dispersal patterns caused by the complex spatial heterogenatiy of the environmental constraints to seed movement. Such unintuitive dispersal patterns may be missed by

Figure 8: Dispersal from several trees with and without a hypothetical highway.

a human analyst, or by models that rely on a simpler representation of the physical environment, such as spatial diffusion models.A third hypothetical scenario involved the simulation of a new highway through the study area. Highway rights-of-way have been cited as facilitators of invasive species dispersal (Lelong *et al.*, 2007; Pauchard and Alaback, 2004; Pauchard and Alaback, 2006), as paved surfaces, combined the maintained vegetation of medians and burms, facilitate seed movement. This simulation was carried out in much the same way as the previous simulation: the friction layer was altered along a hypothetical path to incorporate a 20 meter wide highway-right-of-way with low resistance to seed movement. In addtion to this modification, the elevation layer was altered along the same path to reflect an even grade, which facilitated seed movement from the perspective of both friction and elevation. This had the effect of increasing the dispersal area after secondary dispersal, as showin in figure 8. In this image, the blue line running roughly southwest-to-northeast represents the hypothetical highway, and the colored

points represent dispersal pattern after secondary dispersal. The red points resulted from the unaltered friction surface, and the pink represent those affected by the altered friction and elevation conditions. Note that the strongly linear dispersal pattern of both dispersals is the product of elevation constraints to the northwest and southeast of the trees' starting locations. The primary result of adding in the hypothetical highway was that seeds dispersed farther in all areas surrounding the highway pixels, with particular enhancement of dispersal along the highway, as seen in the northeast of the image.

6. Results

6.1 Comparison of predicted seed dispersal pattern with field data

In order to qualitatively verify the model's prediction accuracy, several field collected datasets were compared to the predicted seed dispersal. The above mentioned primary dispersal dataset collected by Landenberger *et al.* (2007) was used to verify the model's primary dispersal predictions, while datasets collected by McGraw, Kaproth, and Hanna, and McGraw and Kaproth were used for secondary dispersal comparison (Personal communication).

The model's primary dispersal pattern prediction was compared to field data in terms of mean seed distance from origin tree and the density-distance curves. These measurements were made within a GIS, which was used to both measure the average distance of predicted seeds from origin tree and to calculate the seed density at the same transects as defined by the Landenberger *et al.* study. Tables 4 and 5 show that the results are not precisely identical, but are qualitatively very similar. A slight discrepancy can be seen between the model's predicted distance-density curve and that collected by Landenberger *et al.*, in that model predicts slightly lower relative seed density at transects nearer the origin tree, and has a slightly flatter overall shape. This discrepancy is also reflected in the mean distance from origin tree, which is 24.7 meters for the field data and 22.7 meters for the modeled seeds. Overall, the model replicated the field data with reasonable accuracy. However, because the same dataset was used to both parameterize and verify primary dispersal, this verification process illustrates the model's ability to accurately reproduce field collected data, but does not validate the model *per se*, as full validation requires data from other field sites and from other studies.

Table 4: Composite Seed Density from Landenberger *et al.*, (2007)

$y = 0.932e^{-0.03x}$

$R^2 = 0.976$
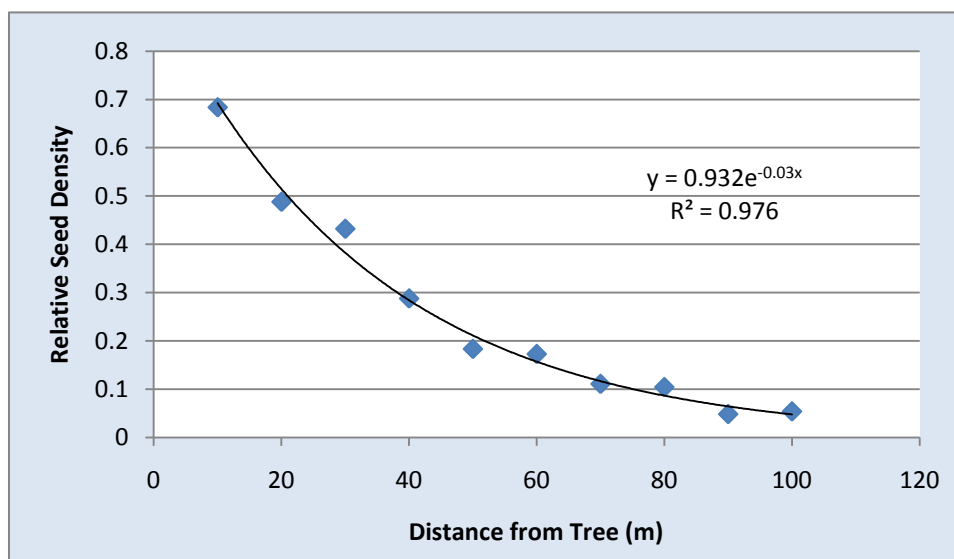
Table 5: Model Predicted Relative Seed Density
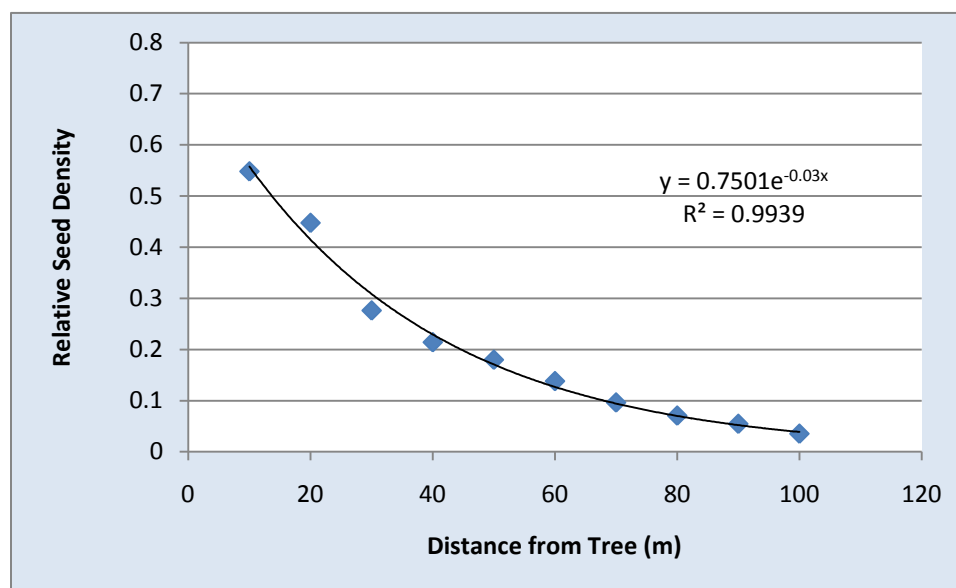
Secondary dispersal data comparison was more qualitative, relying on overall trends from two datasets. The first dataset consisted of primary and secondary dispersal of *A. altissima* seeds along a highway (McGraw, Kaproth, and Hanna,), while the second consisted of secondary dispersal only, collected in a forest setting (McGraw and Kaproth,). The highway dataset indicated a trend for seed dispersal parallel to the highway, which was replicated by the model, as seen in Figure 8 of section 5.5. This figure shows that seeds were predicted to be biased towards movement parallel to highway features, thus corresponding qualitatively to the field data. The second dataset was limited to a forest environment, and was thus applicable only to the grid cells classified as forest in the landcover classification. This dataset indicated that the vast majority of *A. altissima* seeds move one meter or less during secondary dispersal in a forest setting, while a small proportion move up to ten meters. This movement was verified for forested areas within the model by measuring each agent's total secondary movement distance. Again, because the same dataset was used to both parameterize and verify the model, this cannot be taken to be a validation, which would require further field data.

These qualitative comparisons indicate a satisfactory correspondance between field-collected and modelled primary and secondary dispersal. These comparisons are not intended as validations of the model's predictions, but serve only as a general illustration of the model's capacity to replicate real-world seed dispersal patterns.

7. Discussion

7.1 Integration of data types in an ABM environment

The primary goal of this research was to illustrate the utility of ABM for integrating raster and vector data within a vector modeling-environment. Traditional spatial analysis platforms present two obstacles to spatial modeling of ecological processes that require both raster *and* vector datasets and raster *or* vector topology:

1. Data format and methods incompatibility with associated conversion issues.
2. A static model of spatial analysis.

First, interaction between raster and vector datasets is limited due to the fundamentally incompatible nature of these formats.   Whereas the raster data model is constructed as a grid of identically sized and shaped cells, the vector model is based on geometrically defined coordinates and arcs, and is therefore not constrained by discrete cells. As a consequence, cross-format spatial analysis must either incorporate a mechanism for accounting for the data format incongruity or must include a data conversion technique. Typically, conversion is favored, as software platforms tend to be biased in favor of either the raster or vector format, but this conversion inherently introduces an additional degree of uncertainty into the data (Congalton 1997; Hori and Tanigawa, 1993), and constrains the overall analysis to the limits of either one data format or another.

The essential problem of data format conversion lies in a fundamentally different conceptualization of space: the vector model treats space as continuous, or divisible along lines of zero width, whereas the raster model treats space as discrete units of a square tessellation. For example, in order to convert a polygon dataset to raster format, the continuous geometry of each polygons' arcs must be generalized to grid cells of a given size, thus surrendering geometric precision to a certain extent. Also, because the resulting raster dataset is comprised of a uniform rectangular grid, cells along the input polygon's edges will necessarily contain a degree of spurious, artificially generated information, owing to 'rectangularization' of non-rectangle polygon features. The corresponding conversion introduces error in a different way, but with similar data loss and artificial creation. Raster to vector conversion generates vector data whose margins are crenulated at the scale of the original raster's grid cell resolution. This artifact can be removed with a smoothing algorithm, but such a technique is scale-dependent, and may incorrectly smooth, and therefore remove, areas that should be included in the output polygon. This problem exists because of inherent spatial precision limitations of raster datasets; rasters comprise discrete cells, each of which contains a discrete value representing the amalgam of data of the real-life phenomenon in question, extending to the full dimensions of the cell. During the conversion process, it is impossible to determine which area of the cell most characterized the eventual value given to the entire cell, so any smoothing of the grid shape is inherently

presumptive and potentially spurious. Thus, conversion from one data type to the other reduces the predictive reliability of the overall model by introducing a degree of uncertainty. A description of the conceptual basis of the two data formats and their mutual incompatibility issues can be found in Longley *et al.* (2005).

The second major drawback of previous dispersal models is that GIS analysis typically treats vectors (points/lines/polygons) as static, and calculates, for example, distance decay or point clustering by assuming fixed positions. For the application of seed distribution modeling, which entails not only vector/raster display but also the movement of vectors along topologically and environmentally constrained paths, a dynamic model of vectors allows for a more realistic representation of reality.

Agent-based modeling addresses both of these issues. Instead of requiring data to be converted to a single data type, ABM allows both raster and vector datasets to be added to the same modeling environment, thus employing a mechanism of vector/raster interaction not used by traditional analysis. For example, in the seed dispersal model created for this study, the vector-based agents were capable of interacting with raster data by querying the grid cell over which they passed, meanwhile retaining their angular (*i.e.* vector) movement through the environment space. In this way, no data conversion is required, and thus uncertainty due to format conversion is avoided. The second issue is dealt with by the very nature of ABM; agents move across the modeling environment as part of their basic conceptualization, thereby inherently representing a dynamic process.

7.2 Success of ABM in realistically modeling the phenomenon

There are two elements relevant in assessing the success of modeling the physical process of seed dispersal using ABM: 1) the ability to represent the dispersal environment, and 2) the ability to accurately model seed dispersal within this environment.

7.2.1 Representation of dispersal environment

Each data layer used in this model provides information relevant to the dispersal or establishment of seeds. However, each has limitations, some of which can be exacerbated by the ABM platform. These limitations are the product of the spatial and spectral resolutions of each dataset.

Spatial resolution, or the dimensions of each grid cell, determines how precisely any given dataset can control seed movement through space. For this study, the relevant datasets included the NAIP color infrared imagery, the Landsat TM imagery, and the digital elevation model, which, like all raster datasets, are composed of regular grid cells. A dataset of one meter resolution, such as the NAIP color infrared dataset, can affect seed behavior in one square meter increments, while a dataset of three meter resolution, such as the digital elevation model, can

only affect seed moment on a three square meter grid. In reality, the movement of an individual seed is controlled at a much finer spatial scale than remotely sensed data can resolve, and hence any attempt to apply even the finest resolution remote sensing data to seed modeling will be limited. It was assumed that the pixels, while too large to represent the micro-controls on seed movement, could be used to infer the generalized controls of a patch of environment. For example, a pixel classified as 'forest' from the NAIP imagery was assumed to have a generalized control over seed movement for the entire pixel, representing the aggregate control of that patch of environment on seed movement. This generalizes the micro-scale spatial heterogeneity within the pixel and assumes an average control for the entire area. However, these micro-environments within each cell were beyond the purview of this study. Although potentially a limitation, this type of assumption must be made by any pixel-based analysis, and is justified as long as it is made clear that each cell controls seed behavior in purely an average sense.

If the micro-scale controls on seed movement were to be captured directly, the only solution would be to obtain ever finer pixels, until the resolution of each pixel is smaller than the control on the seed. For this study, this would require pixels on the scale of centimeters. Not only is this resolution beyond the capability of modern remote sensing, it would create datasets of enormous file size. Netlogo is incapable, in fact, of incorporating even a one meter resolution dataset covering the entire study area, which is why the imagery was resampled to two meters. A centimeter-resolution dataset would be far beyond the capacity of Netlogo, and would likely require much more powerful software and hardware than is typically available at present.

While spatial resolution limits the precision with which the modeled environment can control seed behavior through space, spectral resolution controls how closely each classified pixel describes the real world. Because spectral resolution applies only to datasets with multiple spectral bands, only the NAIP color infrared and Landsat TM datasets are affected by spectral resolution.

Within a classified image, each informational class, such as landcover type, is statistically characterized and distinguished from every other informational class. For this study, a supervised classification method was used, which defined classes *a priori* using so-called training sites. Because these training sites were user-generated based on a combination of spatial and spectral properties, some classes contain more spectral variation than others. For example, an area defined as low intensity development may contain a mixture of vegetation and man-made structures, and thus the class must incorporate both vegetation and concrete pixels. Classes with increased spectral variation inherently yield increased generalization, which presents a potential limitation for the landuse and landcover datasets used in this study. For example, the 'low intensity development' class in the landuse dataset was used to predict germination likelihood, but since this class is spectrally diverse, the prediction is less powerful. In fact, the spectral complexity of any given class can be used as an indication of its precision in terms of inferring landuse or landcover. Spectrally complex classes represent more a generalized classification of

landuse, for example, and thus potentially lose value in terms of identifying any particular landuse.

Although spectral resolution limits the predictive power of the classified satellite and aerial imagery, these datasets nevertheless represent important sources of information. Classification of complex spectral data yields discrete informational classes, which summarize the imaged landscape in a meaningful way. The informational classes, such as landuse or landcover type, distill the complexity of the environment, allowing the modeler to make judgments about a particular area's control on seed dispersal based not on non-remote-sensing observations, but on *in situ* observations or expert knowledge. In this way, classified imagery can be used to indirectly infer information relevant to seed dispersal or germination.

7.2.2 Seed dispersal

Given that ABM provides an acceptable conceptual modeling space, the obvious challenge in modeling seed dispersal is matching the model's predictions to field measurements. Although field verification and parameterization were not within the purview of this study, the model can incorporate any future parameterization into its existing structure. Thus, the contribution of this model was to create an adaptable framework into which field measurements can be fed, and whose parameters can be finely modified to match empirical data.

The seed dispersal and establishment model created for this study accounted for two basic controls on seed dispersal: species-based primary dispersal traits, and environmental controls on seed movement during secondary dispersal (Chambers and MacMahon, 1994; Schurr *et al.*, 2005). For model simplicity, establishment potential was calculated based on landuse classification alone. In principle, any number of input data layers could be used to define either variable, establishment or dispersal. Movement and establishment controls were parameterized within the model to represent the relative importance of various constraints on seed movement and establishment, such as initial wind dispersal, topographical barriers, surface friction, and landuse type. Each of these variables contributes to the overall pattern of dispersal and establishment, and thus the task is to modify these relative contributions until the dispersal pattern predicted by the ABM is representative of dispersal patterns measured in the field. The ABM environment allows a fairly literal representation of seed dispersal, in that each seed is represented by an individual vector agent, which means that each agent's movements can be tracked, and their interaction with environmental controls on movements investigated.

Primary and secondary dispersal were treated as separate events for this model, so that each could be calibrated individually. The scope of this study included a qualitative field-data comparison for both primary and secondary dispersal. Field data collected by Landenberger *et al.* (2007) were used to calibrate the model's primary dispersal stage by matching the model's primary dispersal output to the *in situ* density-distance curve. Because the model predicted

primary dispersal in 5 meter increments, based on the dispersal curve created by Landenberger *et al.*, it produced a primary dispersal pattern with clearly delimited 5 meter steps. This blocky dispersal is the result of the underlying mechanism for agent movement entailed by the type of available field data. The unnaturally organized spatial pattern of dispersal produced by this mechanism has the key limitation of sharply dividing seed densities into 5 meter zones, meaning that predicted density decreases in a series of steps, rather than continuously. For example, the composite distance-density curve predicted that approximately 12% of all sampled seeds traveled 5 meters and that 10% traveled 10 meters. It was assumed that these proportions included seeds traveling anywhere within the 5 meter buffers of each area; thus 12% fell between 0 and 5 meters, 10% within 5 and 10 meters, and so on. The key limitation of this method is that the model predicts that a seed is approximately 16% more likely to travel 4.99 meters than 5 meters, since 5 meters falls within the 5 to 10 meter zone, while 4.99 falls within the 0 to 5 meter zone. This is clearly an artifact of the model rather than a natural phenomenon.

It was not assumed that this dataset alone endowed the model with perfect predictive accuracy; rather, this dataset was used simply as a representative example of the type of data that could be used to fully parameterize the model. Provided a sufficient body of field data, primary dispersal parameters can be modified within the model to reflect the mean distance, angular direction, and distance-density curve shape.

Because this study dealt specifically with one species of tree, *i.e. A. altissima,* the species-level controls over primary dispersal and establishment potential were standard across the model. However, it is important to note that the model's applicability is not limited to this particular species or a particular dispersal environment. Indeed, one of the key advantages of ABM in this application is that, provided sufficient data, this model could be used to predict the distribution and establishment of any wind-dispersed species.

Secondary dispersal was accounted for by the various surface layers, and was assessed only qualitatively against field data. Although controlled experiments, such as those carried out by Schurr *et al.* (2005), will be required in order to parameterize the control of each environmental layer on secondary seed dispersal of *A. altissima* seeds, this study shows *how* such data is effectively integrated into this type of model. Each layer's contribution to total resistance to seed movement and establishment is independently adjustable, meaning that the combined effect of environmental variables is highly customizable to the specific area of study. In this model, establishment potential is similarly conjectural, and serves only to illustrate the potential of this approach: both secondary dispersal and establishment likelihood are highly tunable based on environmental layer properties, which allows very detailed projections of future spread of *A. altissima* or other similar vegetation.

The primary limitation of the final model created by this study was the lack of field validation. The field measurements to which the modeled predictions were compared were insufficient, both

in terms of sample size and dispersal environment variety, to fully parameterize the model. However, the model's strength lies in its ability to accommodate new data. The model's framework is set up so that new data can be input and used to parameterize the predicted dispersal to match field measurements.

7.3 Practicality of ABM versus conventional invasive species modeling

There are a number of advantages and challenges presented by the proposed agent-based model of seed dispersal, both in terms of data requirements and processing, and in contrast with established invasive species modeling techniques.

Compared to traditional ecological modeling of invasive species and invasibility, the agent-based structure described by this study represents a fundamentally different approach. Modeling invasibility typically involves creation and mapping of ecological niche models, which define the observed and predict the potential distribution extent of a species, based on environmental variables and species' tolerances (Peterson, 2003). Both niche modeling and the proposed agent-based approach define establishment potential based on environmental variables, but whereas the former predicts invasions purely based on environmental characteristics, the latter explicitly models seed transport in addition. This essentially adds a level of predictive ability by defining likely vector routes, while maintaining the environmental controls on establishment and establishment proposed by conventional ecological niche modeling.

While ABM potentially offers increased predictive precision, it also requires a greater amount of data and more computer processing power. For example, creation of a generic ecological niche model may consist of a geographic regression of four or five fairly coarse-scale environmental and climatic data layers (e.g. Brown *et al.,* 2008), which is a relatively straightforward GIS process. In contrast, an agent-based counterpart would require 1) finer-resolution datasets for meaningful seed transport modeling, 2) more substantial preprocessing to weight and parameterize the data layers appropriately, and 3) the locations of existing trees in or near the study area. Thus, the proposed method requires more detailed input data to produce more specific and predictive output predictions. For example, in order to cover the study site, the created ABM was limited to two-meter resolution for the environmental layers, as finer resolution rasters were too large, and caused the program to crash.

Another data-side consideration is consistency of collection and processing methods between study sites. The two imagery datasets used in this model are widely available within the U.S., both spatially and temporally, and conform to precisely defined accuracy and precision standards, and are thus acceptably consistent and available for implementation at various study locations. On the other hand, the lidar data and digital elevation model used present potential data availability concerns. The photogrammetrically created three meter DEM used was generated as part of a state-level mapping and data-generation program, and although USGS data

accuracy standards are well documented, this dataset is available only for selected states and regions. The least widely available dataset employed in this model, however, was the lidar used to create the surface roughness index. The elevation and imagery datasets could potentially be substituted, provided sufficiently similar accuracy and processing, but lidar coverage is constrained in both geographic extent and collection and processing standards. Data availability may thus pose an important constraint to the application of this type of model, and further experimentation is necessary to determine appropriate dataset substitutions, additions, or omissions.

The agent-based model designed by this study was designed to model seed dispersal on a local scale of analysis, and thus produces predictive results that are as consistent between study sites as the datasets themselves, without requiring the structure of the model to be altered on a scene-by-scene basis. However, this modeling approach is not limited to this large geographic scale. In principle, creation of a state or regional model of seed dispersal and establishment is possible via either of two routes. Most simplistically, the study area could be widened by increasing the extent of the input data layers and relying on more powerful computer processing. This extended model would be conceptually identical to the current model, and would eliminate minor inconsistencies and gaps between study sites, thus ensuring consistent predictions across a wider area. However, this would require considerable computing power, and may be beyond the ability of Netlogo, as this program was not designed to handle the large raster datasets entailed by this extent. In fact, it was found that even at the extent of the study area, Netlogo was unable to incorporate rasters of finer than two meters.

Another option to increase the areal coverage would be to aggregate the data layers into larger pixels, thus reducing the memory usage required by the program. This would allow considerably larger areas to be modeled, and could therefore present a more attractive option for modeling the broader invasion patterns of *A. altissima* or another such tree. However, the drawback of this aggregation is that changing the scale of analysis would fundamentally change the modeled controls on seed movement. Surface roughness, for example, can be measured at different spatial scales (Sappington *et al.*, 2007), but its effect on secondary seed dispersal is only relevant when measured at a similar scale to the seeds' movement. Thus, resampling a surface friction layer to a coarse spatial resolution would render it less meaningful for modeling seed dispersal. To accommodate this problem, increased weight would need to be given to macro controls on dispersal, such as aqueous transport (Kaproth and McGraw, 2008), rather than the micro controls accounted for by the current model. This is well within the ability of ABM, but would require further modeling experimentation and construction.

7.4 Potential use for management

The model created by this study was not, and was not intended to be, sufficiently field-verified and parameterized for implementation. However, the model does lay down the groundwork for

such parameterization, possibly as a part of a future study. Provided this fieldwork and fine-tuning, the model could potentially be used as a predicative tool by land or resource managers, or by anyone else with interest in the spread of windborne seeds.   The model created for this study could be used to predict the spread of *A. altissima* and similar trees in two ways. The most straightforward method, and that which formed the basis of the model created for this study, is to assume that all environmental data layers, *i.e.* landcover, elevation, landuse, and surface roughness, are current and static. In this type of model, input of known *A. altissima* locations to the model can be used to predict seed dispersal and establishment, as described throughout this paper. An additional strength of ABM, however, is its ability to incorporate potential changes to the environment in its predictions, thus making seed dispersal predictions based on conjectural future environmental controls on seed movement and establishment.

The hypothetical examples described by figures 7 and 8 of section 5.5 illustrates the utility of the model in predicting the altered pattern of seed dispersal given a change to the dispersal environment. In these examples, the impact of a hypothetical landcover change from grass and trees to pavement is illustrated: the pavement increases the distance traveled by agents that pass over it, therefore increasing the total distribution extent. In this way, various simulations representing a range of changes to the physical environment could be run, each providing specific predictions of changes to seed dispersal in the study area. Example simulations include such scenarios as topographic alteration, urban and suburban expansion, deforestation, and reclamation. *A. altissima* has been shown to spread extensively via highway networks and in and around developing lands (Kowarik and Saumel, 2007). By analyzing the potential impact of proposed new highways or developments, agent-based modeling could be used as a tool to predict the spread of this invasive species, both in terms of seed dispersal and establishment potential.

Because this ABM requires several fairly high-resolution datasets as well as field collected tree locations, data availability poses a potential limitation for this type of modeling. While the imagery datasets are widely available, the limited extent, resolution, and data processing standardization of the elevation and lidar datasets could pose a major constraint to widespread implementation of this ABM as a predictive tool. However, as each year sees the collection of more data, this limitation could be eased in the future. Nevertheless, pending the creation of small-scale, high-resolution lidar and DEM datasets, these data layers pose the biggest drawback to the model created for this study.

8. Conclusions and Future Research

The model created for this study was designed to predict the spread and establishment of airborne seeds across a differentiated environment, using *A. altissima* as a case study. The agent-based environment was found to be suitable for modeling this phenomenon, as it incorporated vector-based movement and raster-based data layers, and because it allowed for precise parameterization of the effect of individual environmental controls on seed movement and establishment. In contrast to ecological-niche-model-based predictions of species spread, the ABM model makes specific, spatially explicit predictions of where seeds are likely to move and germinate, rather than simply defining areas vulnerable to establishment. Although spatial diffusion modeling also accounts for movement through space, it does so using a much more generalized representation of the physical environment, and thus does not account for the effects of a highly heterogeneous dispersal setting. Agent-based modeling allows focused predictions of species spread, accounting for both the complex environment and the process of spatial diffusion. In principle, this model is applicable to any species that employs airborne seed dispersal, provided sufficient parameterization of environmental controls.

The primary limitations of this model are data and computer hardware requirements. Because fairly high-resolution data are required to meaningfully represent environmental controls, this model is limited to fairly data-rich geographic regions. As a result of both the resolution of the environmental data and the number of seeds represented in the model, the model imposes a high demand on current generation desktop computers. This problem limits the predictive precision of the model by limiting the resolution of each dataset, as well as limiting the overall number of datasets in the model.

In addition, computer processing limitations hinder full vector/raster integration. Although the ABM does successfully incorporate vector movement and a raster background, the Netlogo platform cannot incorporate raster datasets that are both of very fine resolution and large extent, thus limiting either the predictive precision or the total study area of the model. This limitation is an artifact of Netlogo's structure, which was designed to focus on mobile agents rather than a complex environment, and not of agent-based modeling generally.

Overall, the ABM environment was found to be an appropriate framework for modeling the phenomenon under study. Further research on field-verified secondary dispersal and establishment potential would dramatically increase the predictive power of this model, and could straightforwardly be integrated into the model to yield a fully implementation-ready model. Extensive field measurement and verification could provide additional avenues for further research, using the ABM as a hypothesis-generation tool. Investigation of the relevance of various environmental controls on seed dispersal and establishment would help to better parameterize this model, both in terms of which layers to include and how much weight to give them. Given an accurately parameterized model of dispersal for a species, the ABM could be

used as a hypothesis generation tool to investigate the spread of seeds in new or modified environments.

9. References

Axtell, R. L., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlund, A. C., Harburger, J., Chakravarty, S., Hammond, R., Parker, J., and Parker, M., 2002. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences,* 99(3): 7275-7279.

Bailey, A. M., Thorne, B. C., and Peirce, S. M., 2007. Multi-cell agent-based simulation of the microvasculature to study the dynamics of circulating inflammatory cell trafficking. *Annals of Biomedical Engineering*, 35(6): 916-936.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(3): 7280-7287.

Brown K. A., Spector, S., and Wu, W., 2008. Multi-scale analysis of species introductions: combining landscape and demographic models to improve management decisions about non-native species. *Journal of Applied Ecology*, 45: 1639-1648.

Burch, P. L. and Zedaker, S. M., 2003. Removing the invasive tree *Ailanthus altissima* and restoring natural cover. *Journal of Arboriculture*, 29(1): 18-24.

Call, L. J., and Nilson, E. T., 2003. Analysis of spatial patterns and spatial association between the invasive tree-of-heaven (*Ailanthus altissima*) and the native black locust (*Robinia pseudoacacia*). *American Midland Naturalist*, 150(1): 1-14.

Congalton, R. G., 1997. Exploring and evaluating the consequenses of rector-to-raster and raster-to-vector conversion. *Photogrammetric Engineering & Remote Sensing,* (63)4: 425-434.

Chambers, J. C. and MacMahon, J. A., 1994. A day in the lLife of a seed: movements and fates of seeds and their implications for natural and managed systems. *Annual Review of Ecological Systems*, 25: 263-292.

Cole, V. and Albrecht, J., 1999. Modelling the spread of invasive species -- parameter estimation using cellular automata in GIS. *Proceedings of the Second International Workshop on Dynamic and Multi-Dimensional GIS.* Beijing: October 4-6, 1999. ISPRS, Beihing National Geomatics Center of China, 32(4W12): 141-148.

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

Davidsson, P., 2002. Agent-based social simulation: a computer science view. *Journal of Artificial Societies and Social Simulation*, 5(1): 1-5.

Earnest D. C., 2008. Coordination in large numbers: An Agent-based model of international negotiations. *International Studies*, 52: 363-382.

Epstein, J, 2002. Modeling civil violence: an agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(3) 7243-7250.

Haltuch, M. A. and Berkman, P. A., 2000. Geographic information system (GIS) analysis of ecosystem invasion: exotic mussels in Lake Erie. *Limnology and Oceanography*, 45(8): 1779-1787.

Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B. A., Moore, K., Taylor, C. and Thomson, D., 2005. The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, 8: 91-101.

Heubner, C. D., 2003. Vulnerability of oak-dominated forest in West Virginia to invasive exotic plants: temporal and spatial patterns of nina exotic species using herbarium records and land classification data. *Castenea*, 68(1): 1-14.

Heywood, V. H., 1989, *Biological Invasions: a Global Perspective*, Drake, J. A., di Castri, F., Groves, R. H., Kruger, F. J., Rejmanek, M., Williamson, M., Mooney, H., A. (Eds.), pp. 31-59 (New York: John Wiley & Sons Ltd) .

Higgins, S. I., Richardson, D. M., and Cowling, R. M., 1996. Modeling invasive plant spread: the Role of Plant-Environment Interactions and Model Structure. *Ecology*, 77(7): 2043-2054.

Hori, O. and Tanigawa, S., 1993. Raster-to-vector conversion by line fitting based on contours and skeletons. *Document Analysis and Recognition*, *Proceedings of the Second International Conference on*, 353-358.

Hu, S. Y. 1979. Ailanthus. *Arnoldia*, 39(2):29-50.

Hurley, A., Watts, D., Burke, B., and Richards, C., 2004. Identifying gypsy moth defoliation in Ohio using Landsat data. *Environmental and Engineering Geoscience*, 10(4): 321-328.

Jennings, N. R., 2000. On agent-based software engineering. *Artificial Intelligence*, 117: 277-296.

Jennings, N. R., 2001. An Agent-based Approach for building complex software systems. *Communications of the ACM*, 44(4): 35-41.

Johnson, B. R., 2009. A self-organizing model for task allocation via frequent task quitting and random walks in the honeybee. *The American Naturalist*, 174(4): 537-547.

Johnson E. A. and Fryer, G. I., 1992. Physical characterization of seed microsites -- movement on the ground. *Journal of Ecology*, 80(4): 823-836.

Joyce, D., Kennison, J., Densmore, O., Guerin, S., Barr, S., Charles, E. and Thompson, N. S., 2006. My way or the highway: a more naturalistic model of altruism tested in an iterative prisoners' dilemma. *Journal of Artificial Societies and Social Simulation, 9(2)*.

Kaproth, M. A. and McGraw, J. B., 2008. Seed viability and dispersal of the wind-dispersed invasive *Ailanthus altissima* in aqueous environments. *Forest Science*, 54(5): 490-496.

Knapp, L. B. and Canham, C. D., 2000. Invasion of an old-growth forest in New York by *Ailanthus altissima*: sapling growth and recruitment in canopy gaps. *Journal of the Torrey Botanical Society*, 127(4) 307-315.

Kota, N. L., Landenberger, R. E., and McGraw, J. B., 2007. Establishment and early growth of *Ailanthus* and tulip poplar in three levels of forest disturbance. *Biological Invasions*, 9: 197-211.

Kowarik, I., 1995. Clonal growth in *Ailanthus altissima* on a natural site in West Virginia. *Journal of Vegetation Science*, 6: 853-856.

Kowarik I., and Saumel, I., 2007. Biological flora of Central Europe: *Ailanthus altissima* (Mill.) Swingle. *Perspectives in Plant Ecology, Evolution and Systematics*, 8(4): 207-237.

Kowarik, I. and Saumel I., 2008. Water dispersal as an additional pathway to invasions by the primarily wind-dispersed tree *Ailanthus altissima*. *Plant Ecology*, 198: 241-252.

Landenberger, R. E., Warner, T. A., and McGraw, J. B., 2009. Spatial patterns of female *Ailanthus altissima* across an urban-to-rural land use gradient. *Urban Ecosystems*, 12(4): 437-448.

Landenberger, R. E., Kota, N. L., and McGraw, J. B., 2007. Seed dispersal of the non-native invasive tree *Ailanthus altissima* into contrasting environments. *Plant Ecology*, 192: 55-70.

Lelong, B., Lavoie, C., Jodoin, Y., and Belzile, F., 2007. Expansion pathways of the exotic common reed (*Phragmites australis*): a historical and genetic analysis. *Diversity and Distributions*, 13: 430-437.

Liebhold, A. M., Halverson, J. A., and Elmes, G. A., 1992. Gypsy moth invasion in North America: a quantitative analysis. *Journal of Biogeography, 19: 513-520*.

Liebhold, A. M. and Tobin, P. C., 2008. Population ccology of insect invasions and their management. *Annual Review of Entomology*, 53: 387-408.

Lodge, D. M., and Shrader-Freschette, K., 2002. Nonindigenous species: ecological explanation, environmental ethics, and public policy. *Conservation Biology,* 17(1): 31-37.

Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., 2005. *Geographic Information Systems and Science*, 2nd Edition. (Wiley).

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., and Bazzaz, F. A., 2000. Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, 10(2): 689-710.

Merchant, J. W. and Narumalani, S., 2009. Integrating remote sensing and geographic information systems. In Warner, T., Nellis, M. D., and Foody, G. M. (eds.) *The SAGE Handbook of Remote Sensing*, 257-268. (Sage Publications, London, UK).

Merriam, R. W., 2003. The abundance, distribution and edge associations of six non-indigenous, harmful plants across North Carolina. *Journal of the Torrey Botanical Society*, 130(4): 283-291.

Miller, N. P. and Matlack, G. R., 2010. Population expansion in an invasive grass, *Microstegium vimineum*: a test of the channeled diffusion model. *Diversity and Distributions*, 16: 816-826.

Parks, S. C., Garifullin, M., and Dronzek, R., 2005. Argus invasive species spread model constructed using agent-based modeling approach and cellular automata. *Proceedings of the 2005 Winter Simulation Conference*, Orlando, Fl, USA, December 4-7. http://www.informs-sim.org/wsc05papers/prog05.html, data last accessed June 25, 2009.

Pauchard, A. and Alaback, P. B., 2004. Influence of elevation, land use, and landscape context on patterns of alien plant invasions along roadsides in protected areas of south-central Chile. *Conservation Bioilogy*, 18(1): 238-248.

Pauchard, A. and Alaback, P. B., 2006. Edge type defines alien plant species invasions along *Pinus contorta* burned, highway and clearcut forest edges. *Forest Ecology and Management*, 223: 327-335.

Peterson, A. T., 2001. Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, 103(3): 599-605.

Peterson, A. T., 2003. Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology,* 78(4): 419-433.

Pimentel, D., Lach, L., Zuniga, R., and Morrison, D., 2000. Environmental and economic costs of nonindigenous species in the United States. *BioScience*, 50(1): 53-65.

Reaney, S. M., 2008. The use of agent-based modeling techniques in hydrology: determining the spatial and temporal origin of the channel flow in semi-arid catchments. *Earth Surface Processes and Landforms*, 33: 317-327.

Sala, O. E., Chapin III, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, N., Poff, N. L., Sykes, M. T., Walker, B. H., Walker, M., and Wall, D. H., 2000. Global biodiversity scenarios for the year 2100. *Science*, 287: 1770-1774.

Sappington, J. M., Longshore, K. M., and Thompson, D. B., 2007. Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *The Journal of Wildlife Management*, 71(5): 1420-1426.

Schurr, F. M., Bond, W. J., Midgley, G. F. and Higgins, S. I., 2005. A mechanistic model for secondary seed dispersal by wind and its experimental validation. *Journal of Ecology*, 93(5): 1017-1028.

Spears, B. M., Dull, C. W., Rubel, D. N., 1991. Use of geographic information system in gypsy moth integrated pest management. *Resource Technology* 90. Proc. Second International Symposium on Advanced Technology in Natural Resources Management, Washington, DC, 1990, ASPRS.

Sycara, K, P., 1998. Multiagent systems. *The American Association for Artificial Intelligence*, 0738-4602-1998: 79-92.

Vitousek, P. M., D'Antonio, C. M., Loope, L. L., and Westbrooks, R., 1996. Biological invasions as global environmental change. *American Scientist*, 84: 468-478.

Webster, C. R., Jenkins, M. A., and Jose, S., 2006. Woody invaders and the challenges they pose to forest ecosystems in the eastern United States. *Forest Ecology*, 104(7): 366-374.

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., and Losos, E., 1998. Threats to imperiled species in the United States. *BioScience*, 48: 607-615.

Wilensky, U., 1999. Netlogo. *Center for Connected Learning and Computer-Based Modeling.* Northwestern University, Evanston, IL.

Woldridge, M., 1997. Agent-based software engineering. *IEEE Proceedings of Software Engineering*, 144(1): 26-37.

10. Appendix: Netlogo Code

```
extensions [gis]
globals [friction-dataset
        ailanthus-dataset
        germination-dataset
        elevation
        slope
        aspect]

patches-own [friction grow]
turtles-own [energy moved]

breed [firstWave]
breed [secondWave]
breed [thirdWave]
breed [fourthWave]
breed [fifthWave]
breed [stopped]
breed [established]
breed [downstream]

to export
  gis:store-dataset gis:turtle-dataset turtles user-new-file
end

to setup-turtles
  clear-turtles
  set ailanthus-dataset gis:load-dataset (word "finalDatasets_rev3/ailanthus/" ailanthusLayer
".shp")
 setup-firstwave
 setup-secondwave
 setup-thirdwave
 setup-fourthwave
 setup-fifthwave
end

to setup-patches
set elevation gis:load-dataset (word "finalDatasets_rev3/grids/" elevationLayer ".asc")
set friction-dataset gis:load-dataset (word "finalDatasets_rev3/grids/" frictionLayer ".asc")
gis:set-world-envelope gis:envelope-of  friction-dataset
let min-friction gis:minimum-of friction-dataset
let max-friction gis:maximum-of friction-dataset
gis:apply-raster friction-dataset friction
set germination-dataset gis:load-dataset "finalDatasets_rev3/grids/germ2.asc"
gis:apply-raster germination-dataset grow
let horizontal-gradient gis:convolve elevation 3 3 [ 1 1 1 0 0 0 -1 -1 -1 ] 1 1
let vertical-gradient gis:convolve elevation 3 3 [ 1 0 -1 1 0 -1 1 0 -1 ] 1 1
set slope gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
set aspect gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
 let x 0
  repeat (gis:width-of slope)
  [ let y 0
    repeat (gis:height-of slope)
    [ let gx gis:raster-value horizontal-gradient x y
      let gy gis:raster-value vertical-gradient x y
      gis:set-raster-value slope x y sqrt ((gx * gx) + (gy * gy))
      ifelse (gx != 0) or (gy != 0)
      [ gis:set-raster-value aspect x y atan gy gx ]
      [ gis:set-raster-value aspect x y 0 ]
```

```
      set y y + 1 ]
    set x x + 1 ]
  gis:set-sampling-method aspect "bilinear"

ask patches
 [
   if (friction <= 0) or (friction >= 0)
   [set pcolor scale-color black friction min-friction max-friction ] ]

end

to setup-firstwave
  foreach gis:feature-list-of ailanthus-dataset
  [foreach gis:vertex-lists-of ?
    [foreach ?
      [let location gis:location-of ?
        if not empty? location
          [ create-firstWave 100000
            [ set xcor item 0 location
              set ycor item 1 location
              set HEADING 0 set color red]
          ]]]]
end

to setup-secondwave
  foreach gis:feature-list-of ailanthus-dataset
  [foreach gis:vertex-lists-of ?
    [foreach ?
      [let location gis:location-of ?
        if not empty? location
          [ create-secondwave 300
            [ set xcor item 0 location
              set ycor item 1 location
              set HEADING 0 set color white ]
          ]]]]
end

to setup-thirdwave
  foreach gis:feature-list-of ailanthus-dataset
  [foreach gis:vertex-lists-of ?
    [foreach ?
      [let location gis:location-of ?
        if not empty? location
          [ create-thirdwave 300
            [ set xcor item 0 location
              set ycor item 1 location
              set HEADING 0 set color orange ]
          ]]]]
end

to setup-fourthwave
  foreach gis:feature-list-of ailanthus-dataset
  [foreach gis:vertex-lists-of ?
    [foreach ?
      [let location gis:location-of ?
        if not empty? location
          [ create-fourthwave 300
            [ set xcor item 0 location
              set ycor item 1 location
              set HEADING 0 set color blue ]
          ]]]]
end
```

```
to setup-fifthwave
  foreach gis:feature-list-of ailanthus-dataset
  [foreach gis:vertex-lists-of ?
    [foreach ?
      [let location gis:location-of ?
        if not empty? location
          [ create-fifthwave 300
            [ set xcor item 0 location
              set ycor item 1 location
              set HEADING 0 set color yellow ]
          ]]]]
end

to primary-dispersal
  move-firstWave
  move-secondWave
  move-thirdWave
  move-fourthWave
  move-fifthWave
  setup-secondary
end

to setup-secondary
  ask stopped
    [ set energy 40 set heading 0 set color green ]
end

to secondary-dispersal
 move-stopped

end

to move-firstwave    ;;NOTE: This section is repeated for all five waves, but is only copied here
                     ;;for the first.

  ask firstwave
  [ right random-normal 360 360]
  ask firstwave
 [
   let a random-float 1
   ifelse (a < .1357 )
                                     [let b (0)
                                     let c gis:raster-sample elevation patch-ahead b
                                     let d gis:raster-sample elevation self
                                     let ee (c - d)
                                     ifelse ee <= 15
                                       [ forward b set breed stopped]
                                       [forward b - (b / 80)]
                                       ][

     ifelse ( a > 0.135734046 and a < 0.253091686948606 )
                                     [let bb (random-float 1)
                                     repeat random 1[
                                     let cc gis:raster-sample elevation patch-ahead 1
                                     let dd gis:raster-sample elevation self
                                     let eee (cc - dd)
                                     ifelse eee <= 1
                                       [ forward 1 set breed stopped]
                                       [forward bb - (bb / 80)]]][
```

```
ifelse ( a > 0.253091686948606 and a < 0.354560819642024)
                                  [let bb (random-float 1)
                                  repeat random 2[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
                                    [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.354560819642024 and a < 0.442292515653038)
                                  [let bb (random-float 1)
                                  repeat random 3[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
                                    [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.442292515653038 and a < 0.518146623012215)
                                  [let bb (random-float 1)
                                  repeat random 4[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
                                    [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.518146623012215 and a < 0.583731193525193)
                                  [let bb (random-float 1)
                                  repeat random 5[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
                                    [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.583731193525193 and a < 0.640436572299271)
                                  [let bb (random-float 1)
                                  repeat random 6[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
                                    [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.640436572299271 and a < 0.689464872046625)
                                  [let bb (random-float 1)
                                  repeat random 7[
                                  let cc gis:raster-sample elevation patch-ahead 1
                                  let dd gis:raster-sample elevation self
                                  let eee (cc - dd)
                                  ifelse eee <= 1
```

```
                                        [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

  ifelse ( a > 0.689464872046625 and a < 0.731855456997892)
                                        [let bb (random-float 1)
                                        repeat random 8[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

  ifelse ( a > 0.768506976666542 and a < 0.800196416563768)
                                        [let bb (random-float 1)
                                        repeat random 9[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.800196416563768 and a < 0.827595569725182 )
                                        [let bb (random-float 1)
                                        repeat random 10[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.827595569725182 and a < 0.851285278233672 )
                                        [let bb (random-float 1)
                                         repeat random 11[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.851285278233672 and a < 0.871767746648337 )
                                        [let bb (random-float 1)
                                        repeat random 12[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

  ifelse ( a > 0.871767746648337 and a < 0.889477188375212)
                                        [let bb (random-float 1)
                                         repeat random 13[
```

```
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][

ifelse ( a > 0.889477188375212 and a < 0.904789030675089 )
                                    [let bb (random-float 1)
                                    repeat random 14[
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][

ifelse ( a > 0.904789030675089 and a < 0.918027873447898 )
                                    [let bb (random-float 1)
                                    repeat random 15[
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][

ifelse ( a > 0.918027873447898 and a < 0.929474370514065 )
                                    [let bb (random-float 1)
                                    repeat random 16[
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][

ifelse ( a > 0.929474370514065 and a < 0.939371179271011 )
                                    [let bb (random-float 1)
                                    repeat random 17[
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][

ifelse ( a > 0.939371179271011 and a < 0.947928104853167 )
                                    [let bb (random-float 1)
                                    repeat random 18[
                                    let cc gis:raster-sample elevation patch-ahead 1
                                    let dd gis:raster-sample elevation self
                                    let eee (cc - dd)
                                    ifelse eee <= 1
                                      [ forward bb set breed stopped]

                                      [forward bb - (bb / 80)]]]][
```

```
ifelse ( a > 0.947928104853167 and a < 0.955326547847952 )
                              [let bb (random-float 1)
                               repeat random 19[
                               let cc gis:raster-sample elevation patch-ahead 1
                               let dd gis:raster-sample elevation self
                               let eee (cc - dd)
                               ifelse eee <= 1
                                 [ forward bb set breed stopped]

                                  [forward bb - (bb / 80)]]]][

ifelse ( a > 0.955326547847952 and a < 0.961723349856052 )
                              [let bb (random-float 1)
                               repeat random 20[
                               let cc gis:raster-sample elevation patch-ahead 1
                               let dd gis:raster-sample elevation self
                               let eee (cc - dd)
                               ifelse eee <= 1
                                 [ forward bb set breed stopped]

                                  [forward bb - (bb / 8)]]]][

ifelse ( a > 0.961723349856052 and a < 0.967254118419068 )
                              [let bb (random-float 1)
                               repeat random 21[
                               let cc gis:raster-sample elevation patch-ahead 1
                               let dd gis:raster-sample elevation self
                               let eee (cc - dd)
                               ifelse eee <= 1
                                 [ forward bb set breed stopped]

                                  [forward bb - (bb / 80)]]]][

ifelse ( a > 0.967254118419068 and a < 0.972036101800594 )
                              [let bb (random-float 1)
                               repeat random 22[
                               let cc gis:raster-sample elevation patch-ahead 1
                               let dd gis:raster-sample elevation self
                               let eee (cc - dd)
                               ifelse eee <= 1
                                 [ forward bb set breed stopped]

                                  [forward bb - (bb / 80)]]]][

ifelse ( a > 0.972036101800594 and a < 0.976170674563971 )
                              [let bb (random-float 1)
                               repeat random 23[
                               let cc gis:raster-sample elevation patch-ahead 1
                               let dd gis:raster-sample elevation self
                               let eee (cc - dd)
                               ifelse eee <= 1
                                 [ forward bb set breed stopped]

                                  [forward bb - (bb / 80)]]]][

ifelse ( a > 0.976170674563971 and a < 0.979745486639167 )
                              [let bb (random-float 1)
                               repeat random 24[
                               let cc gis:raster-sample elevation patch-ahead 1
```

```
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.979745486639167 and a < 0.982836321437422 )
                                        [let bb (random-float 1)
                                        repeat random 25[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.982836321437422 and a < 0.985508702404352 )
                                        [let bb (random-float 1)
                                        repeat random 26[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.985508702404352 and a < 0.98781928206925 )
                                        [let bb (random-float 1)
                                        repeat random 27[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.98781928206925 and a < 0.989817043037432)
                                        [let bb (random-float 1)
                                        repeat random 28[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

ifelse ( a > 0.989817043037432 and a < 0.991544336385772 )
                                        [let bb (random-float 1)
                                        repeat random 29[
                                        let cc gis:raster-sample elevation patch-ahead 1
                                        let dd gis:raster-sample elevation self
                                        let eee (cc - dd)
                                        ifelse eee <= 1
                                          [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][
```

```
ifelse ( a > 0.991544336385772 and a < 0.993037779474679 )
                                [let bb (random-float 1)
                                repeat random 30[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
                                  [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.993037779474679 and a < 0.994329032209455 )
                                [let bb (random-float 1)
                                repeat random 31[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
                                  [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.994329032209455 and a < 0.995445468207219 )
                                [let bb (random-float 1)
                                repeat random 32[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
                                  [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.995445468207219 and a < 0.996410755097642 )
                                [let bb (random-float 1)
                                repeat random 33[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
                                  [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.996410755097642  and a < 0.997245356259454 )
                                [let bb (random-float 1)
                                repeat random 34[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
                                  [ forward bb set breed stopped]

                                    [forward bb - (bb / 80)]]]][

ifelse ( a > 0.997245356259454 and a < 0.997966964629162 )
                                [let bb (random-float 1)
                                repeat random 35[
                                let cc gis:raster-sample elevation patch-ahead 1
                                let dd gis:raster-sample elevation self
                                let eee (cc - dd)
                                ifelse eee <= 1
```

```
                                          [ forward bb set breed stopped]

                                            [forward bb - (bb / 80)]]]][

    ifelse ( a > 0.997966964629162 and a < 0.99859087777841 )
                                      [let bb (random-float 1)
                                      repeat random 36[
                                      let cc gis:raster-sample elevation patch-ahead 1
                                      let dd gis:raster-sample elevation self
                                      let eee (cc - dd)
                                      ifelse eee <= 1
                                        [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

    ifelse ( a > 0.99859087777841 and a < 0.999130322211354 )
                                      [let bb (random-float 1)
                                      repeat random 37[
                                      let cc gis:raster-sample elevation patch-ahead 1
                                      let dd gis:raster-sample elevation self
                                      let eee (cc - dd)
                                      ifelse eee <= 1
                                        [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][
    ifelse ( a > 0.999130322211354 and a < 0.999596733756911 )
                                      [let bb (random-float 1)
                                      repeat random 38[
                                      let cc gis:raster-sample elevation patch-ahead 1
                                      let dd gis:raster-sample elevation self
                                      let eee (cc - dd)
                                      ifelse eee <= 1
                                        [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

    ifelse ( a > 0.999596733756911 and a < 1 )
                                      [let bb (random-float 1)
                                      repeat random 39[
                                      let cc gis:raster-sample elevation patch-ahead 1
                                      let dd gis:raster-sample elevation self
                                      let eee (cc - dd)
                                      ifelse eee <= 1
                                        [ forward bb set breed stopped]

                                          [forward bb - (bb / 80)]]]][

]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]
  ask firstwave [set breed stopped]
end

to move-stopped

  ask stopped
  [ let x random-normal 7 .8
    let y gis:raster-sample germination-dataset self
    ifelse x < y
      [set breed established set color brown]
      [right random-normal 360 360
          let m random-exponential 2
          let i gis:raster-sample elevation patch-ahead m
          let h gis:raster-sample elevation self
```

```
            let j (i - h)
            ifelse (j <= 1)
             [
            ifelse friction <= 1  [ forward random m + 5 set energy energy - 1 ] [
            ifelse friction = 2  [ forward random m + 4.5 set energy energy - 2 ] [
            ifelse friction = 3  [ forward random m + 4 set energy energy - 3 ] [
            ifelse friction = 4  [ forward random m + 3.5 set energy energy - 4 ] [
            ifelse friction = 5  [ forward random m + 3 set energy energy - 5 ] [
            ifelse friction = 6  [ forward random m + 2.5 set energy energy - 6 ] [
            ifelse friction = 7  [ forward random m + 2 set energy energy - 7 ] [
            ifelse friction = 8  [ forward random m + 1.5 set energy energy - 8 ] [
            ifelse friction = 9  [ forward random m + 1 set energy energy - 9 ] [
            ifelse friction = 10 [ forward random m + 0.5 set energy energy - 10 ] [
            ifelse friction = 11 [ forward random m + 0 set energy energy - 11 ] [
            ifelse friction = 12 [ forward random m set energy energy - 12 ] [
            ifelse friction = 13 [ forward 0 set energy energy - 40 ] [
]]]]]]]]]]]]]][ forward 0]]]]

end
to display-elevation
 clear-all
  set elevation gis:load-dataset "finalDatasets_rev3/grids/dem.asc"
  gis:set-world-envelope gis:envelope-of elevation
  let horizontal-gradient gis:convolve elevation 3 3 [ 1 1 1 0 0 0 -1 -1 -1 ] 1 1
  let vertical-gradient gis:convolve elevation 3 3 [ 1 0 -1 1 0 -1 1 0 -1 ] 1 1
  set slope gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
  set aspect gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
  let x 0
  repeat (gis:width-of slope)
  [ let y 0
    repeat (gis:height-of slope)
    [ let gx gis:raster-value horizontal-gradient x y
      let gy gis:raster-value vertical-gradient x y
      gis:set-raster-value slope x y sqrt ((gx * gx) + (gy * gy))
      ifelse (gx != 0) or (gy != 0)
      [ gis:set-raster-value aspect x y atan gy gx ]
      [ gis:set-raster-value aspect x y 0 ]
      set y y + 1 ]
    set x x + 1 ]
  gis:set-sampling-method aspect "bilinear"
  gis:paint elevation
end
```