Individual-based modeling and visualization of forest succession and biomass within the Prospect Hill tract of Harvard Forest in Petersham, Massachusetts.

Joshua Sims

ABSTRACT

The carbon storage and dynamics of terrestrial forest vegetation will play a major role in determining the degree of effects of climate change. As the response to climate change increases, the mapping and modeling of forest carbon stocks and dynamics will become of increasing importance at scales ranging from the individual forest stand, to the landscape level and ultimately global forest inventories. Individual-Based Modeling offers a way to investigate the above ground biomass distribution within a forest stand based upon differential ecological and historical phenomena which had led to a specific distribution of individual trees. This analysis will help to inform us of both how different model parameters affect biomass distribution and the influence of changing model parameters on model outcomes. We have parameterized and evaluated the SORTIE model of forest succession to reflect the effects of spatially explicit historical events upon the current biomass distribution across ten Prospect Hill plots at the Harvard Forest, MA, USA. From data recorded during the NASA DESDynI field campaign (Cook, 2010), biomass in each of these plots were estimated using the Jenkins set of allometric equations (Jenkins, 2003). Each initialization simulates a history of one particular plot. The most important parameters to calibrate were the species specific Asymptotic Diameter Growth and Slope of Diameter Response Curve. To further this investigation, we have used the output from these model runs to produce three dimensional representations of the forest structure within these plots. This methodology can be used to improve biomass remote sensing techniques. Sensitivity analysis using Monte Carlo methods to test the effects of randomly perturbing the yearly growth rate and intrinsic mortality probability parameters shows the model to be functioning properly. Future

areas of research include investigating the nature of the path dependency inferred by natural and anthropogenic disturbance and more realistic methods of creating three dimensional forest structures.

Contents

1. INTRODUCTION 8

1.1 Scientific Rationale 8

1.2 Objectives 10

1.3 Individual-Based Models 11

2. STUDY SITE 13

3. DATA AND MODELS 15

3.1 Harvard Forest GIS Data 15

3.2 Harvard Forest Field Campaign Data 19

3.3 The SORTIE Model 20

4. ANALYSIS METHODS 21

4.1 Objective 1: History of Prospect Hill Plots and Model Initialization 21

4.1.1 Prospect Hill Plot 1 23

4.1.2 Prospect Hill Plot 2 23

4.1.3 Prospect Hill Plot 3 24

4.1.4 Prospect Hill Plot 4 25

4.1.5 Prospect Hill Plot 5 25

4.1.6 Prospect Hill Plot 6 26

4.1.7 Prospect Hill Plot 7 26

4.1.8 Prospect Hill Plot 8 27

4.1.9 Prospect Hill Plot 9 27

4.1.10 Prospect Hill Plot 10 28

4.2 Objective 2: Model Parameterization 29

4.2.1 Parameter estimation 30

4.2.2 Biomass calculations 33

4.3 Objective 3: Model Performance Evaluation 34

4.3.1 Application of model to Field Plots outside Prospect Hill 34

4.3.2 Data processing 35

4.3.3 Probability Density Functions 35

4.3.3 Statistical Testing 35

4.3.4 Sensitivity Analysis 36

4.4 Objective: Visualizations 37

5. RESULTS 38

5.1 History of Prospect Hill Plots and Model Initialization 38

5.2 Modeling 41

5.3 Evaluation of Model Performance 42

5.3.1 Plots outside Prospect Hill 42

5.3.2 Probability Density Functions 43

5.3.3 Statistical Testing 44

5.3.4 Sensitivity Analysis 45

5.4 Visualizations 47

6. DISCUSSION 48

7. CONCLUSIONS 61

8. REFERENCES 63

9. APPENDIX 1: ADDITIONAL FIGURES 67

10. APPENDIX 2: SORTIE PROCESSING SUITE CODE 79

I would like to express thanks to the following individuals for their contributions:

Kathleen Bergen:

Associate Professor, University of Michigan School of Natural Resources

Paul Siqueira:

Associate Professor, University of Massachusetts Amherst

William Currie:

Associate Professor, University of Michigan School of Natural Resources

Steven Pacala:

Director, Princeton Environmental Institute and Professor, Princeton Department of Ecology and Evolutionary Biology

Shannon Brines:

Lab Manager, University of Michigan School of Natural Resources Environmental Spatial Analysis Laboratory

1. INTRODUCTION

1.1 Scientific Rationale

The carbon storage and dynamics of terrestrial forest vegetation will play a major role in the mitigating the effects of increasing atmospheric carbon dioxide on global temperature rise. However, the extent and present and future dynamics of forest carbon are often poorly understood due to the lack of accurate carbon inventories (Fearnside 2000), and the uncertainty of tree responses to elevated levels of carbon dioxide (Zak et al 2007). As the response to climate change increases, the mapping and modeling of forest carbon stocks and dynamics will become of increasing importance at scales from landscapes to regions to globally. Reliable estimates are necessary (Fearnside 1997), however field measurement techniques are intensive in both cost and time and rely on statistical samples. Most field procedures to inventory carbon over areal extents currently involve intensive, stratified ground plot sampling to estimate aboveground dry biomass (Brown 1999).

In order to more explicitly map the spatial distribution of forest biomass carbon and change in carbon due to disturbance or succession, remote sensing methods are increasingly employed. Lidar, (LIght Detection And Ranging) remote sensing is of particular interest due to its ability to directly measure forest height and density, parameters that are understood to have quantifiable relationships with forest biomass (Nilsson 1986). While maps of biomass and its spatial variation derived from lidar have been suggested as very useful, remote sensing data alone cannot typically be used for purposes such as quantifying the effects of climate change over a region or predicting the effects of different disturbances regimes on future forest composition and biomass carbon.

The above highlights the need for the development of accurate models of the carbon balance and dynamics within forest ecosystems. Parameterization of a complex model of forest succession with both field data and historical records should allow for more realistic predictions about both the current and future distribution of carbon within such forest sites. Proper calibration of models should allow for scenario-based extrapolation of the fate of biomass and carbon into the future. Inclusion of historical disturbance events will allow for even better projections of the effects of management and disturbance into the future under the effects of climate change and invasive pests. Models that allow testing the effects of past disturbances and management strategies on the current distribution of biomass within a given site can lead to future management decisions to maximize carbon sequestration.

Models that have spatial capabilities built into them will be most useful in coupling with remote sensing data, such as that from lidar (Sun and Ranson, 2000). Although a number of potential models exist, an additional issue with models is that they are typically developed and parameterized for a particular area and the effects of modifying that parameterization for use in a novel location may be unknown, even if for the same general ecoregion. Here we have two research questions: Can we use an existing model with spatial capabilities to create output that would be useful for anticipated remote sensing data such as that from lidar, and how well does that model perform in terms of biomass estimates and canopy structure? If we use an existing forest model that was developed elsewhere and for which we may need to modify parameters, can we develop a method to estimate whether our parameter choices were scientifically sound?

1.2 Objectives

The Harvard Forest Long Term Ecological Research site (LTER) has one of the most comprehensively recorded ecological histories of any Northeastern USA Forest. The goal of this study is to combine this rich ecological record with an individual based model (IBM) of forest dynamics to create a realistic model of the biomass distribution of ten research plots within the Prospect Hill Tract of the Harvard Forest and then to assess the success of those models.

There are four specific objectives. The first objective is to create an ecological history for each plot by combining maps of time of agricultural abandonment, 1938 hurricane damage, silviculture treatments and current composition and field-measured biomass. The second objective is to parameterize an individual-based model of forest succession to simulate the forest dynamics and biomass within each plot from abandonment to its current state. The third objective is to evaluate the parameterization of the model in two ways: first to compare our results against current field data for each of the ten calibration plots as well as three reserved plots not used for model development area; second to assess proper model functioning through a Monte Carlo simulation-based sensitivity analysis of model parameters. The final fourth objective is to use these data to create a three dimensional model of the forest using realistic canopy structures and the spatial capabilities of the SORTIE model.

1.3 Individual-Based Models

The spatial complexity of individual-based forest models has evolved with advancements in computing power. Theoretically, individual-based forest models attempt to capture system dynamics though tracking the collective interactions of individual tree stems, each possessing certain traits. In early models such as JABOWA (Botkin et al 1972) and FORET (Shugart, 1984; Shugart & West, 1977), the forest is simulated as collections of individuals cohabiting gap-sized patches of land. The tree crowns are represented as a horizontal plane at the top of each tree, but tree locations within gaps are not specified. . Early gap models did not include interactions between patches. In 1974 the FOREST model added spatial complexity by providing some information on vertical structure; however interactions among trees still focuses on horizontal interactions. (Ek & Monserud 1974).

Derived from FORET, ZELIG tracks the locations and interactions between the gaps within a given forest stand (Smith Urban 1988). Yet the spatial location of individual trees is still not specified, and height is the only tree crown dimension parameter. The consideration of interaction between plots allowed the modeling of large-gap disturbances and the resulting increase in shade-intolerant trees, an emergent behavior which FORET is unable to produce (Urban et al 1991). SPACE achieved higher spatial resolution by attributing Cartesian coordinates to individual trees in a stand (Bushing 1991).

The approach of SPACE was extended with horizontal and vertical structure modeling in DRYADES to simulate boreal coniferous forests of British Columbia

(Mailly et al 2000). This approach allowed understory trees to persist within proximity to large over story trees. Pacala et al (1993) employed a similar approach in modeling a Connecticut forest. Hurtt et al (2010) used lidar measured canopy heights to initialize the Ecosystem Demographic (ED) model developed by Moorcroft (2001) to predict biomass. SORTIE, the model we use in this study, also added a significantly more complex light regime, with seasonal and daily movements of the sun used to calculate available light. More complex algorithms replaced the regeneration, growth and mortality processes of the simpler gap models. SORTIE has a number of useful spatial and 3D features. Within SORTIE, each individual tree is attributed a Cartesian coordinate, and the canopy has a height, width and depth.

In forests, the physical propagation, extinction, and scattering of light is in part a function of the depth, width and structure of forest tree crowns and gaps between them. Spatially explicit 3D simulated forests from IBMs could be coupled with a lidar scattering model. First a 3D scene of the forest plot could be constructed and divided into cubic cells of a size corresponding to the resolution of the lidar system. The resulting cells are attributed to the different structural components of the forest. A physics-based scattering algorithm can be used to estimate the lidar backscatter from each of these cells (e.g. Sun and Ranson, 2000). For both landscape to regional carbon modeling and linkage with lidar remote sensing, a coupled model approach which: a) could be used to predict biomass and b) is based on more ecologically robust and spatially explicit 3D forest models, would be an important next step forward. The aim of this study is to develop a spatially explicit model to support such a remote sensing approach.

2. STUDY SITE

The Harvard Forest Long Term Ecological Research site (LTER) is a 1200 hectare (ha) experimental forest acquired by Harvard University in 1907 (Foster, 1992). Located near Petersham, Massachusetts, the forest is broken in three tracts, named Prospect Hill, Tom Swamp and Slab City (Figure 1). The general native forest type is *Transitional Hardwoods* of central New England (Stehman et al 2003). Most of the forest is composed of artificially planted stands over abandoned agricultural fields (Fisher 1921).

Prospect Hill is the 375 ha northernmost tract within Harvard Forest. The tract is located at roughly 42°32'N and 72°10'W and is situated at an average elevation of 340 meters above sea level. The soil is generally an acidic sandy loam with variable drainage. The dominant forest cover is Red-maple dominated mixed hardwood, with an older growth hemlock-dominated section in the center of the tract. Seven percent of the tract was planted as red-pine (*Pinus resinosa* Ait.*)* plantations by the Civilian Conservation Corps in the 1930s. Historically, the land use included cultivated crops, improved pastureland, unimproved pastureland and permanent woodlot (Foster, 1992). The average annual temperature of Prospect hill is 8.5 degrees Celsius and receives an annual average precipitation of 105 cm with 150 cm of snowfall. (Rasche, 1953)

Tom Swamp is a 475 ha tract to the south of Prospect Hill. The tract surrounds a reservoir, and exhibits much wetter soil conditions than Prospect Hill, with an average elevation of 240 m. The predominant cover type is hemlock dominated conifer stands. Slab City is the southernmost 200 ha tract of Harvard Forest. The dominant forest cover is red-maple and red-oak dominated hardwood forest. The mean elevation is 265 m.

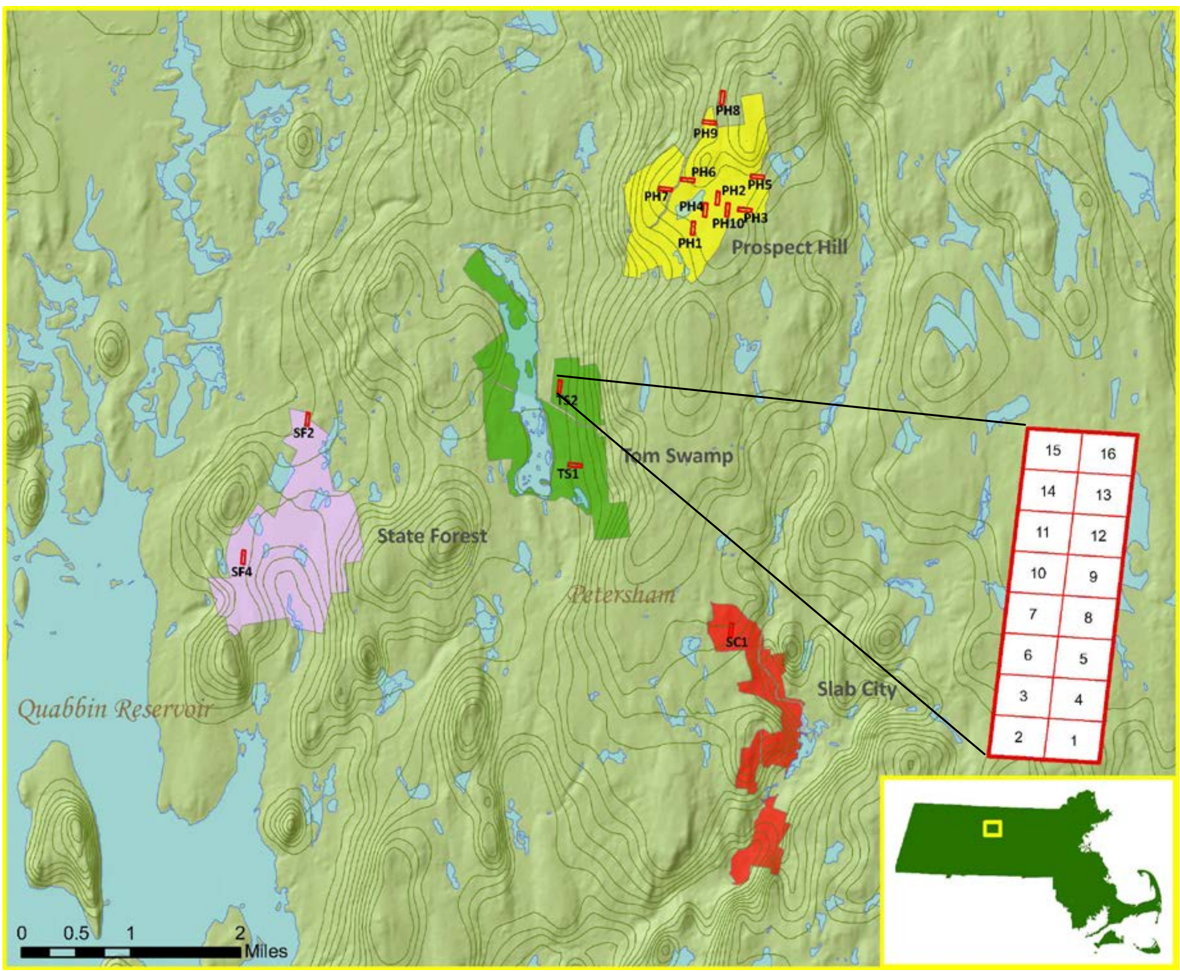


Figure 1. Harvard Forest tracts and plots. The inset shows a 1-hectare plot with its sixteen 25m by 25m subplots numbered one through sixteen.

3. DATA AND MODELS

This section describes the input datasets and the modeling software used in the analysis. The land-use history spatial data was compiled from the Harvard Forest Data Archive and further processed at the University of Michigan School of Natural Resources and Environment (SNRE) Environmental Spatial Analysis Laboratory (ESALab). Field data was collected in 2009 as part of the NASA Deformation, Ecosystem Structure and Dynamics of Ice (DESDynI) project (Cook et al, 2011). The SORTIE-ND model of forest succession was developed by Steven Pacala and associates (Pacala 1996) and further parameterized by us for this study.

3.1 Harvard Forest GIS Data

The first layer used in constructing a GIS map of Prospect Hill is a historical land-use map (*landuse_history.shp*). This layer documents the historical land use within Prospect Hill based upon historical records compiled by David Foster. The land-use types are cultivated crops, mowing improved, unimproved pasture and woodlot (Foster and Boose 1999, Figure 2)

Because some study plots had previously been converted to agriculture and then abandoned, a layer on agricultural extent and abandonment was selected (*PH_agricultural_abandonment.shp*) for model initialization purposes. This layer provides recorded dates of agricultural field abandonment within the Prospect Hill tract of Harvard Forest (Foster and Boose, 1999).

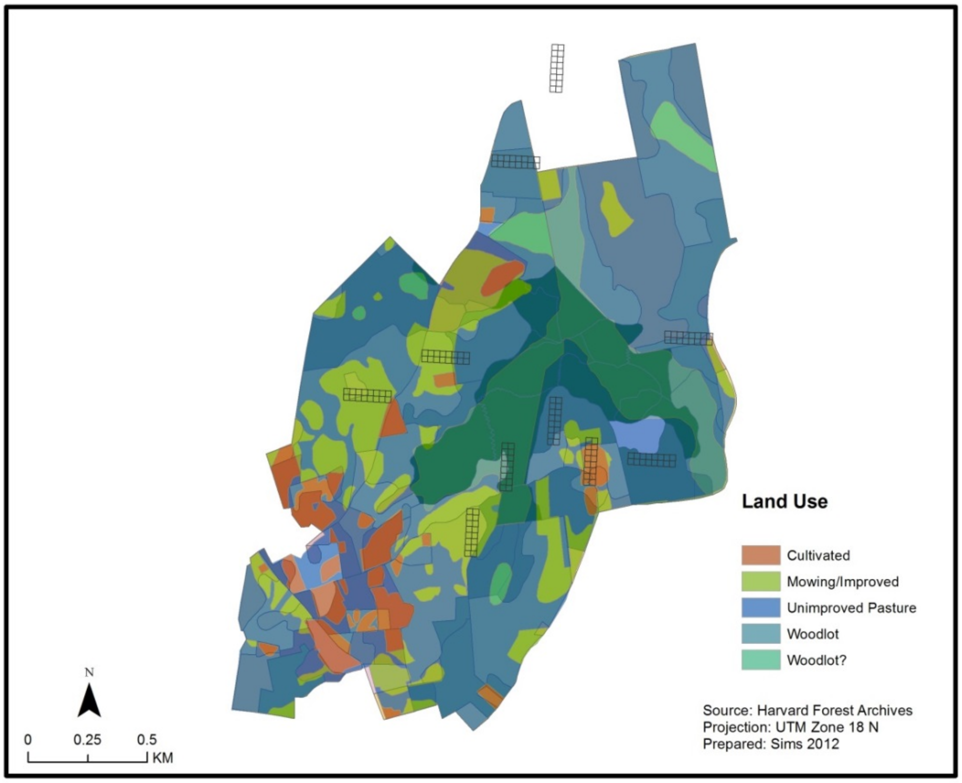


Figure 2. Prospect Hill Historical Land Use. Compiled from historical land parcel records by David Foster (1992) Prospect Hill Plots are overlayed, PH 8 falls outside the coverage of the map.

The first fields were abandoned in the 1840s, with abandonment continuing up through the 1980s. The layer also provides information on areas that had not been converted to agriculture, for example, the central portion of the Prospect Hill tract is characterized as a permanent woodlot (Figure 3). Field plot PH8 is not included within the coverage of this layer.

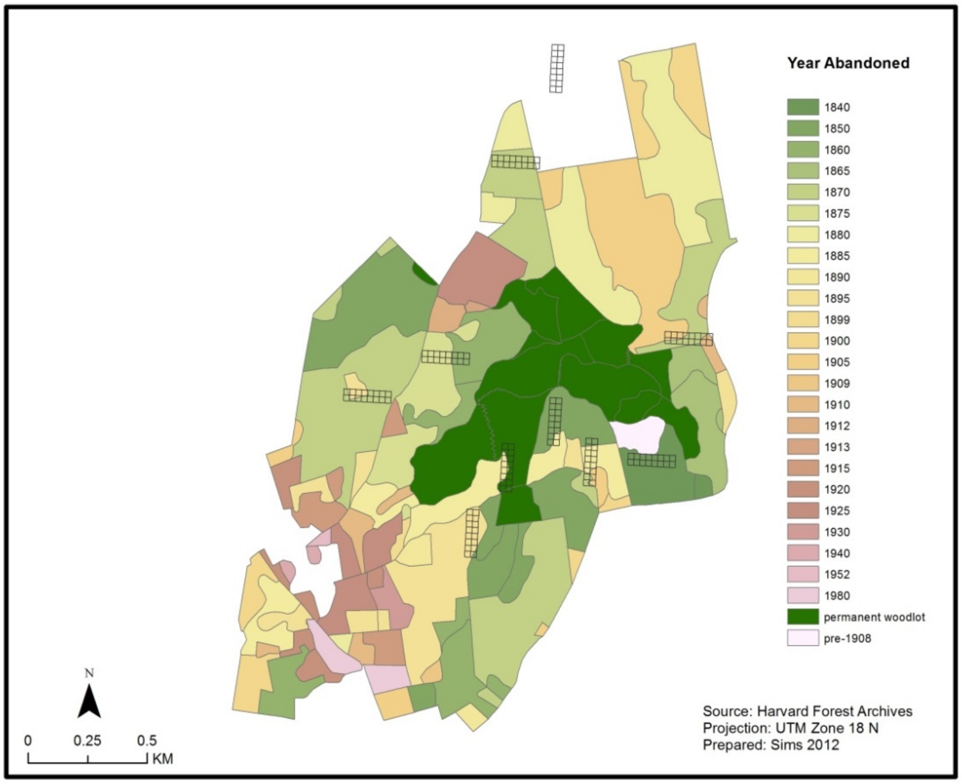


Figure 3. Prospect Hill Agricultural Abandonment. This map displays the year parcels of land within the Prospect Hill tract of Harvard Forest were abandoned from agricultural use. Prospect Hill Plots are overlaid, PH8 lies outside the coverage of the map.

Included in the Harvard Forest GIS is a layer (*1938_hurricane_damage.shp*) which displays the damage to each stand as percentage of all trees in each stand felled by a hurricane which impacted the area in 1938 (Hall, 2005). The data was collected by Will Rowland between 1938 and 1939 (Roland, 1939). The degree of damage is expressed as a percentage and represents the proportion of all dominant and co-dominant trees which were uprooted, leaning or broken off (Hall, 2005; Figure 4)**.** Directly after the hurricane all the downed trees were salvaged, and in 1949 the most plots were thinned to control the dense regeneration after the hurricane. Field plots PH8 and PH9 were not included within the coverage of this layer.

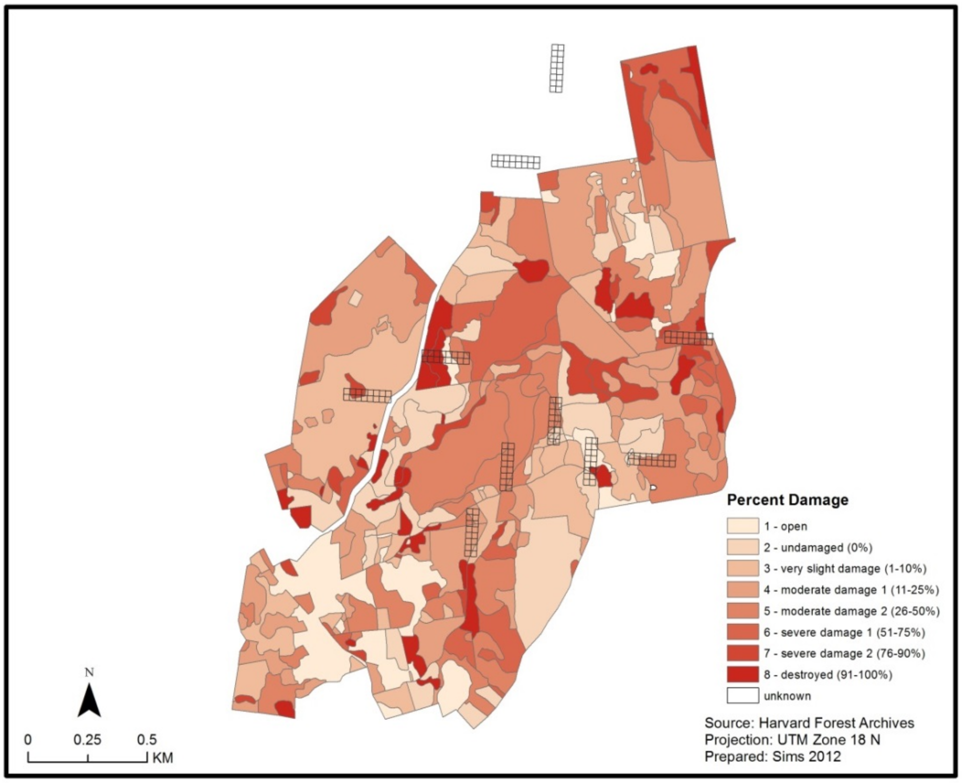


Figure 4. Prospect Hill 1938 Hurricane Damage. Damage is expressed as the percent of dominant or co-dominant trees felled by during the 1938 New England hurricane. Prospect Hill Plots are overlayed, PH8 and PH9 fall outside the coverage of the map.

A fourth GIS layer (*silviculture_shp*) tracks all management activities or silvicultural treatments within the forest from 1909 until 1994 (Hall, 2005). As treatments overlap spatially and temporally, the polygons representing them also overlap. In order to determine all treatments applied at a given site the silviculture layer was intersected with a polygon representing the area of interest to extract the overlapping silviculture treatments. The resulting intersected polygons were output to a database to develop a silviculture history for each area of interest (Hall, 2005).

3.2 Harvard Forest Field Campaign Data

Forest composition, structure and biomass ground validation plots were established and surveyed during the summer of 2009. These field data were subsequently transcribed into electronic format at Goddard Space Flight Center (GSFC) and stored in Oak Ridge national Laboratory Distributed Active Archive Center (Cook et al 2011). Ten plots were established at the Prospect Hill tract. Additionally, two plots were measured in the nearby Harvard Forest Tom Swamp tract and one in the Slab City tract. Each plot was one ha, measuring 200m by 50m. The location of each plot by the sampling teams took into consideration the forest species composition, age/structure, topography, and lidar coverage. The one-ha plots were divided into 16 subplots, each measuring 25m by 25m. All plots locations were established using GPS and these coordinates were used to create a GIS layer of plot and subplot boundaries (ESALab 2010).

Data collected in the field for all plots/subplots included the diameter (dBH, diameter at breast height 1.37 m above the ground), species and condition of each individual tree larger than 10 cm. Canopy height was collected for a subset of dominant trees in each subplot. Statistics calculated from this field data were done at the individual tree, subplot and plot levels. In addition to the data directly collected in the field (e.g. species, stem diameter), basal area and biomass were calculated for individual stems. Biomass was calculated for each stem using the allometric equations of Jenkins et al

(2004). At the plot levels, species proportional composition, stem density/ha, mean dBH, basal area/ha, and total above ground biomass/ha were compiled.

3.3 The SORTIE Model

SORTIE (or SORTIE-ND) is a forest model simulator based on individual tree competition for forest resources, principally light. It is a descendent of the JABOWA-FORET family of gap-based forest models. SORTIE has the ability to predict large-scale forest dynamics from a model parameterized exclusively from individual trees. The model accomplishes this by using the positions of tree crowns in three-dimensional space along with a complex routine used to calculate light regimes beneath forest canopies. SORTIE does not take into account the soil and drainage characteristics of the site. The light available to each tree is then used to calculate species-specific growth rate and risk of mortality. Surviving trees produce seedlings as an increasing function of tree size, and the seedlings are dispersed away from the parent tree in accordance with estimated dispersal functions. The model uses a time step of 5 years. Forest dynamics and simulated tree locations emerge as the collective result of these localized interactions among trees. (Pacala et al 1993).

These simulations can then be used to derive bulk forest parameters such as diameter distributions, height distributions, biomass/carbon/carbon value, basal area and species composition. External disturbance events such as harvest, planting, storms, insect infestations and other mortality events can be included as behaviors in the model. SORTIE was developed using empirical data derived from the Green Mountain Forest study site in northeast Connecticut. (Pacala et al 1993) The current version SORTIE-ND has been parameterized for nine tree species of the Northeastern Forests. The Prospect Hill ground validation plots contain all of these trees; however *Pinus resinosa,* Ait. *(*red pine), which dominates three of the Prospect Hill stands, had not to date been included in the SORTIE model.

4. ANALYSIS METHODS

4.1 Objective 1: History of Prospect Hill Plots and Model Initialization

The first objective was to create an ecological history for each plot using maps from the time of agricultural abandonment, 1938 hurricane damage and silviculture treatments. The results of this were used to initialize a SORTIE model for each Prospect Hill Plot. The year of abandonment was determined from Foster's Land-use history paper and the Prospect Hill Tract GIS in the Harvard Forest Archive (Foster, 1992; Hall, 2005). The agricultural abandonment, land-use history, silviculture treatments, and 1938 hurricane damage shapefiles from the Prospect Hill Tract Harvard Forest Properties GIS were combined Harvard Forest history and disturbance GIS (Foster, 1992; Hall, 2005). .

In order to determine the important disturbance regime in the Prospect Hill plots, the silviculture treatments, natural disturbance and 1938 hurricane damage maps were intersected with the Prospect Hill DESDynI Plot polygons in the Harvard Forest History and Disturbance GIS. The resulting polygons from intersection indicated within each plot the resulting percent damage to the dominant or co-dominant trees during the hurricane, the year of agricultural abandonment and all silvicuture treatments applied throughout time.

These data were used as inputs to a discrete SORTIE model for all ten plots. Each model was initialized by translating the vector GIS data onto the raster grids used by each corresponding SORTIE behavior. While the SORTIE simulations are conducted on the one-ha scale, the scale of interest in this study is 25 x 25 meters. The cell size of the SORITE grids is 10 x 10 meters, so the results from the GIS analysis of each subplot were translated onto the scale of the grids. The year of agricultural abandonment was used as time step zero (T_0) If there was differential abandonment within a plot, the entire plot was initialized using the oldest abandonment age, and a clear cut applied in each time step to the cells which were still in agricultural use. This method models the maintenance clearing of woody plants from agricultural fields. The hurricane damage was modeled using a partial harvest, as the canopy trees were salvaged immediately after the hurricane. The silviculture treatments were implemented as either a harvest or planting based upon the historical records. The DESDynI field data indicates red maple *Acer rubrum* L., which has a relative dominance greater than 25 percent in six of the Prospect hill plots. (Cook et al 2011). According to Motzkin et. al., historical land use is the strongest indicator for the dominance of red maple, which is most competitive on agricultural land uses, especially former unimproved pastures and woodlots. (Motzkin et al., 1999) For this reason the initial sapling densities, which are set to 1 per hectare for all nine species, were modified to model this dominance, as SORTIE does not take into account soil characteristics affected by agriculture. The initial density of red maple was increased to 25 seedlings per hectare.

4.1.1 Prospect Hill Plot 1

Prospect Hill Plot 1 represents the first of two *Pinus resinosa* Ait. (red pine) plantation plots within the field campaign. The Prospect Hill GIS agricultural abandonment layer indicates the majority of this plot was abandoned from agriculture in 1895 (Foster and Boose 1999). Thus, as SORTIE uses five year time steps, the simulations were started 23 time steps ago, with T_0 representing 1895 and T_{23} representing 2010. The 1938 hurricane layer indicated differential damage across the plot (Hall, 2005), which was simulated during time step nine, by implementing the resulting tree damage as reduction in basal area corresponding to the GIS layer (1 – 10%, 11 – 25% and 26 – 50%). Silviculture records indicate the plot was clear cut in 1923 and then planted as a Red pine plantation, however a triangle encompassing most of subplots 4 and 5 was planted with *Picea glauca* Voss (white spruce) and *Larix laricina* Koch (tamarack) (Hall 2005), which are not important in the current forest and therefore not modeled in this study. This historical management regime was modeled with a clear cut and a planting of red pines in a grid pattern with 1.5 m spacing to achieve the tree density surveyed in 2009, with the exception area planted with spruce and Tamarack. The hurricane damage seems to mirror the abandonment, with the portion of the plot abandoned in 1870 receiving minor (10%) damage, while the subplots abandoned in 1900 received severe (80%) damage (Hall, 2005).

4.1.2 Prospect Hill Plot 2

Prospect Hill Plot 2 is one of the two ‘late successional’ *Tsuga canadensis* L. (Eastern hemlock) dominated stands within the field campaign plots within this specific tract. The

Prospect Hill GIS indicated different times of agricultural abandonment in Prospect Hill plot 2. The northern portion of the plot was abandoned in 1850, while subplots 1 and 4 were abandoned in 1885 (Foster and Boose 2005). This was simulated in SORTIE by starting the simulations with T_0=1850 and T_{32} = 2010. The thirty five years during which subplot 1 and 4 were still under agricultural use were simulated by implementing a clear-cut in these subplots for the first seven time steps, and the forest starting to grow in T_7. Silviculture records indicate the subplots 1, 4, 13 and 16 were clear cut in 1925-26, and this was simulated using the harvest function in T_{15} (Hall 2005). The clear cuts led to these subplots escaping damage in the 1938 hurricane (Hall 2005), thus, the storm damage implemented in year T_{18} was not applied to these plots. Subplots 7, 8, 9, 10, 11, 12, 14 and 15 incurred 50% damage during this time step, while subplots 2, 3, 5 and 6 only received slight (10%) damage (Hall 2005). Silviculture records also indicate a thinning in subplots 7, 8, 9, 10, 12, 14 and 15 in 1957, which was simulated using a 30% harvest (Hall 2005).

4.1.3 Prospect Hill Plot 3

Prospect Hill Plot 3 represents one of the red maple dominated plots within this field campaign. The Prospect Hill GIS indicated the entire Prospect Hill Plot PH3 was abandoned in 1850 (Foster and Boose 1999). Thus, the SORTIE simulations began with T_0=1850 and T_{32} =2010. The 1938 Hurricane damage layer indicated the Southwest portion of the plot incurred 50% damage, while the Northwest corner of the plot only sustained minor damage (~10%) (Hall 2005). These values were used to simulate the hurricane damage in T_{18}. Silviculture records also indicated a thinning in 1949 (Hall 2005), which was modeled as a 40% partial harvest in year T_{20}.

4.1.4 Prospect Hill Plot 4

Prospect Hill Plot 4 represents the oldest growth, late successional hemlock dominated plot within the field campaign. The Prospect Hill GIS indicates most of Prospect Hill Plot 4 fell within the 'permanent woodlot' section of the Prospect Hill Tract (Foster and Boose 1999). The records also indicate a clear cut in 1780 (Hall 2005). Thus the simulation was initiated in 1730, with a clear cut occurring in 1780. The section which was abandoned in 1880 (Foster and Boose 1999) was simulated by implementing clear-cuts each time step in subplots 6, 7, 10 and 11 until the year 1880. The GIS layer for the 1938 hurricane indicates homogenous moderate damage (26-50%) across the plot (Hall 2005), which was modeled using 40 percent damage in the corresponding time step.

4.1.5 Prospect Hill Plot 5

Prospect Hill Plot 5 is currently dominated by a mixture of upland hardwoods, mainly red maple and *Quercus rubra* L. (red oak). The Prospect Hill GIS indicated the middle section of Prospect Hill Plot 5 was abandoned in 1870 (Foster and Boose 1999). This was simulated as T_0, with clear cuts occurring in the northwest subplots (10, 11, 14, and 15) until T_6=1900. Subplots 1, 2, part of 3 and 4 were abandoned in T_8=1910 (Foster and Boose 1999), which was simulated with clear cuts up until the eighth time step. The 1938 Hurricane damage layer indicates the subplots 3 and 6-16 sustained moderate damage (26-50%)(Hall 2005), which was modeled using 40% storm damage in T_{13}. The lower portions of subplots 4, 5 and 8 received severe damage (76-90%) (Hall 2005) which was modeled using 80% storm damage. Silviculture records indicate a thinning in 1949 to control the rapid growth after the hurricane damage (Hall 2005). This was modeled with a partial harvest of 50% in $T_{16.}$

4.1.6 Prospect Hill Plot 6

Prospect Hill Plot 6 is representative of the Red oak dominated forest of the Prospect Hill tract. The Prospect Hill GIS indicates the western section of Prospect Hill Plot 6 (subplots 1-19) was abandoned in 1875, while the eastern section (subplots 11-16) was abandoned in 1860 (Foster and Boose 1999). Thus, the eastern section simulation began with T_0=1860, with a clear cut occurring during the first three time steps in western section. Silviculture records indicate a saw timber cutting in the middle section of the plot in 1937, leaving this section open for the hurricane occurring in 1938 (Hall 2005). The section to the west of this cut was completely destroyed (Hall 2005), which was modeled with a clear cut in $T_{16.}$ For modeling purposes, these sections were both clear cut at the same time. The portion of the plot to the east of the cutting received moderate damage (26-50%) (Hall 2005), which was modeled using 40% storm damage. The Natural Disturbance layer of the Prospect Hill GIS indicates a fire affected the plot in 1957 (Foster and Boose 1999). As SORTIE does not include a fire disturbance behavior in its model, here a partial harvest of 70% was applied. This behavior was calibrated to reflect the current tree density and small average diameter within the plot.

4.1.7 Prospect Hill Plot 7

Prospect Hill Plot 7 is one of the red pine plantations surveyed in this study. The Prospect Hill GIS indicates the majority of the plot was abandoned from agriculture in 1870, while the subplots 10, 11, 13 and 14 were abandoned in 1900 (Foster and Boose 1999). This was modeled starting the simulations with T_0 = 1870 and T_{28} = 2010. Clear cuts were implemented in each time step until T_6=1900 in subplots 10, 11, 14 and 14. Silviculture records indicate the plot was clear cut in 1923 and then planted as a red pine plantation (Hall 2005). This was modeled with a clear cut and a planting of red pines in a grid pattern with 1.5 m spacing to achieve the tree density surveyed in 2009. The hurricane damage seems to mirror the abandonment, with the portion of the plot abandoned in 1870 receiving minor (10%) damage, while the subplots abandoned in 1900 receiving severe (80%) damage (Hall 2005).

4.1.8 Prospect Hill Plot 8

Prospect Hill Plot 8 is almost completely dominated by red maple. However, the plot falls outside the coverage of the Prospect Hill and Harvard Forest GIS datasets (Foster and Boose 1999; Hall 2005). Thus, the year of abandonment and effects of the hurricane were estimated from both the closest polygon and most similar field campaign plot, PH3.

4.1.9 Prospect Hill Plot 9

Prospect Hill Plot 9 is currently dominated by red maple, with oak and *Betula alleghaniensis* Britton (yellow birch) as co-dominants. The Prospect Hill GIS indicates Prospect Hill Plot 9 was abandoned in 1870 (Foster and Boose 1999). Although the easternmost portion of the plot falls outside the map layer, for modeling purposes uniform abandonment is assumed. The model begins with T_0 1870 and ends with T_{28} = 2010. Silviculture records indicate a uniform “cutting” in 1910 (Hall 2005). This is modeled by implementing a selective partial harvest of 40% in T_8. Records also indicate a partial harvest and improved cut in 1958 (Hall 2005), which is modeled as another 40% selective partial harvest in T_{18}. Prospect Hill Plot 9 falls outside the coverage of the 1938 Hurricane Damage layer (Hall 2005), so the model uses the damage from the most proximate polygon included in the layer, moderate damage (50%).

4.1.10 Prospect Hill Plot 10

The Prospect Hill GIS indicates a very complex history for Prospect Hill Plot 10. This complex history has led to great variance in aboveground biomass at the 25 by 25m subplot level and currently this plot is dominated by Red maple. Subplots 15 and 16 were abandoned from agriculture in 1850, while subplots 5 and 7-17 were abandoned in 1890. Subplots 1 and 4 were abandoned in 1870 and subplots 2, 3, and 6 were abandoned in 1905 (Foster and Boose 1999). This differential abandonment scheme is modeled starting with T_0=1850 and running through T_{32}=2010. The subplots abandoned after 1850 have a clear cut applied every time step until the corresponding time step of abandonment indicated in the historical record. The 1938 Hurricane Damage map also indicates a complex pattern of damage to PH10 (Hall 2005). Silviculture records indicate a clear cut in subplots 7-16 in 1936, the year before the hurricane damaged the rest of the plot (Hall 2005). However, subplot 6 was completely destroyed during the hurricane. Uncannily, the portion of the plot to the south remained undamaged (Hall 2005). This heterogeneous damage shows the complexity of effects upon the forest of large infrequent disturbances such as hurricanes. The effects of this can be seen in the variance in density and biomass across the plot in both the simulations and the field campaign measurements.

4.2 Objective 2: Model Parameterization

The second objective is to parameterize an SORTIE to simulate the forest dynamics within each plot from abandonment to its current state for the species measured in the Harvard Forest field plots (Cook 2010, Table 1). The initialized SORTIE grids for each Prospect Hill plot were run first using the initial SORTIE parameters (Table 2).

Acronym	**Latin Name**	**Common Name**
ACRU	*Acer rubrum* L	Red maple
ASCA	*Acer saccharum* Marshall	Sugar maple
BEAL	*Betula alleghaniensis* Britton	Yellow birch
FAGR	*Fagus grandifolia* Ehrh.	American beech
TSCA	*Tsuga Canadensis* L.	Eastern hemlock
FRAM	*Fraxinus americana* L.	White ash
PIST	*Pinus strobus* L	White pine
PRSE	*Prunus serotina* Ehrh.	Black cherry
QURU	*Quercus rubra* L	Northern red oak
PIRE	*Pinus resinosa* Aiton	Red pine

Table 1. Species modeled in the Prospect Hill SORTIE Simulations.

Species	**Asymptotic Diameter Growth (A)**	**Slope of Diameter Growth Response (S)**
ACRU	0.167	0.270
ASCA	0.125	0.159
BEAL	0.169	0.137
FAGR	0.152	0.075
TSCA	0.229	0.051
FRAM	0.226	0.025
PIST	0.230	0.019
PRSE	.0249	0.064
QURU	.0266	0.022

Table 2. SORTIE initial growth parameters

4.2.1 Parameter estimation

In forest stands, the density and biomass of trees is determined by a complex interplay between ecological factors which control 1) growth, 2) mortality, 3) species specific growth response, and 4) senescence and dispersal parameters. The underlying ecological factors controlling the above four parameters change between sites due to different soil conditions, hydrologic conditions and elevation; this is why these factors were chosen for modification to model the Prospect Hill research plots. The canopy light penetration parameters, which were empirically derived by Pacala et al (2003), do not change by location with similar species thus the initial SORTIE parameters were used.

Using initial parameters provided along with the SORTIE program, preliminary simulations produced plots with unrealistically high biomass and too few trees with average DBH much higher than the field measurements. To accurately model the forest dynamics of Prospect Hill, it was necessary to adjust the SORTIE growth, mortality senescence and dispersal parameters. All parameters adjustments fall within the prior distributions included with the SORTIE documentation (Pacala 1996, 2010). The initial tree growth parameter provided with the SORTIE-ND program is 0.3 cm per year. While this growth rate produced realistic results in the Connecticut forest in which the SORTIE model was developed, in the Prospect Hill simulations this growth rate produced forest stands with a mean dBH much higher than the field measurements. Through systematically decreasing the growth rates and using a two-sample t-test to compare the modeled results to 2009 ground truth data, a reasonable annual Prospect Hill growth rate was determined to have a mean of 0.1 cm per year with a standard deviation of 0.02 cm per year (Cook et al 2011).

The Adult Background Mortality Rate included in the original SORTIE program is 0.01, which means any given tree has a one per cent chance of dying in any given year. In the Prospect Hill simulations it seemed too few adult trees were dying, growing over-large and dominating the forest stands. Systematically increasing the mortality rate using a two sample t-test comparing simulated to field densities used to generate a mean Adult Background Mortality Rate of 0.018 with a standard deviation of 0.005 (Cook et al 2011).

Originally, the species specific growth parameters (Asymptotic Diameter Growth and Slope of Diameter Growth Response) produced forests dominated with a late successional mix of Hemlock and Beech trees. Only two Prospect Hill plots are dominated by Hemlock, and none of the ten plots have a significant relative dominance of beech. It was necessary to adjust the species specific growth parameters to produce forest stands with similar composition to the Prospect Hill ground validation plots (Tables 2 and 3). As previously mentioned the DESDynI field data indicates red maple *Acer rubrum* L., which has a relative dominance greater than 25 percent in six of the Prospect hill plots. (Cook et al 2011). According to Motzkin et. al., historical land use is the strongest indicator for the dominance of red maple, which is most competitive on agricultural land uses, especially former unimproved pastures and woodlots. (Motzkin et al. 1999) To model this dominance in succession, the Asymptotic Diameter Growth was increased from the initial value of 0.156 to 0.2 and the Slope Diameter Response was increased from 0.0015 to 0.0024. The rest of the Asymptotic Diameter Growth and Slope Diameter Response values were altered to reflect each species current relative dominance in the field plots (Cook et al 2011) All the changed species specific growth parameters were within their prior distributions according to Steven Pacala (Pacala, 2010).

The Standardized Total Recruits (STR) parameter provided with the initial SORTIE package was 0.09 seedlings per year. This parameter controls the number of seedling recruits each adult tree produces each year. In tracking the tree life stage distribution of the simulations it was found far too few seedlings were being produced, which exacerbated the problem of over large old trees dominating the simulated forest stands. Increasing the STR parameter to 1.0 increased the number of seedlings to a reasonable amount of seedlings per hectare as counted in the field measurements.

In order to add red pine into the SORTIE simulations, the species parameters had to be estimated. Most parameters were based upon those generated by Pacala et al (1996) of white pine; however the maximum height and crown radius parameters were reduced to reflect field measurements taken at the University of Michigan Saginaw Forest, a north temperate forest of similar latitude and forest species composition.

Species	Asymptotic Diameter Growth (A)	Slope of Diameter Growth Response (S)
ACRU	0.200	0.240
ASCA	0.00*	0.00*
BEAL	0.155	0.160
FAGR	0.152	0.050
TSCA	0.229	0.051
FRAM	0.180	0.015
PIST	0.16	0.014
PRSE	0.16	0.019
QURU	0.18	0.018
PIRE	0.2395	0.0415

Table 3. Prospect Hill SORTIE modified growth parameters. *ACRU was not significant in any plot.

4.2.2 Biomass calculations

Detailed output files from a SORTIE run record data for each individual tree at each time step. Tree aboveground biomass was calculated within the SORTIE program using the allometric equations generated by Jenkins et al. (Jenkins et all 2004). The Jenkins study aimed to develop generalizable equations by compiling allometric equation studies from all over the continental United States. A species specific set of allometric coefficients along with a tree's diameter is entered into the general power equation:

$$M = aD^b$$

Where M is the above ground biomass of a tree, D is the diameter at breast height and *a* and *b* are species specific allometric coefficients (Table 4).

Species	**Equation Id**	**Correction Factor**	**a**	**b**	**C**	**d**	**DBH units**	**Biomass units**
ACRU	2	1.01	-2.01074	0	2.363	1	Cm	kg
ASCA	2	1.016	-2.034	0	2.451	1	Cm	kg
BEAL	2	1.02	-2.1306	0	2.451	1	Cm	kg
FAGR	1	1	2.1112	2.462	1	0	Cm	g
TSCA	2	1	0.8645	0	2.3859	1	In	lb
FRAM	1	1.105	1.27844	1.4248	2	0	Cm	g
PIST	2	1	5.281	0	2.3069	1	Cm	g
PRSE	1	1.017	1.1981	1.5876	2	0	Cm	g
QURU	2	1	4.9667	0	2.394	1	Cm	g
PIRE	2	1	5.281	0	2.0369	1	Cm	g

Equation 1: $log(biomass) = a + b * (log(dia^c))$
Equation 2: $ln (biomass) = a + b * dia + c * (ln(dia^d))$

Table 4. Allometric parameters. Jenkins equations, parameters and coefficients. All final biomass values were calculated in units of Mg/ha.

Some equations contained in the Jenkins database specify a correction factor to correct for bias introduced by using log functions. The parameters published in the

Jenkins database come from many sources which use a range of units. The SORTIE program performs all unit conversions and reports final biomass in metric tons (Mg).

4.3 Objective 3: Model Performance Evaluation

In the third objective the objective is to evaluate the parameterization of the model both against current field data for each plot and applied to plots outside the research area. A second evaluation step uses Monte Carlo simulations to conduct a sensitivity analysis to investigate and ensure proper functioning of the model.

4.3.1 Application of model to Field Plots outside Prospect Hill

To test the validity of this parameterization of the SORTIE model over areas outside the model development plots, the parameters were applied to three plots also at Harvard Forest but outside of the Prospect Hill Tract. These plots were Tom Swamp Plot 1, Tom Swamp Plot 2 and Slab City Plot 1. While falling within the coverage of the 1938 Hurricane Damage and Silviculture Treatments GIS layers, there is no information of the year of agricultural abandonment for these plots. The hurricane damage and silviculture treatments were programmed into the SORTIE interface as with the Prospect Hill Plots, however the age of the plots had to be estimated based upon the average DBH of the trees and the degree of succession of the forest type. Thus, the choice of time step 0 is considered somewhat imprecise and a potential source of error. The number of time steps of the plot within Prospect Hill most like the validation plot was used.

4.3.2 Data processing

In the detailed output file each SORTIE, the x-coordinate, y-coordinate, dbh, height, canopy height and biomass for each individual tree were recorded. These trees were then attributed to their proper 25 x 25 meter subplot using the *sortie_read* algorithm we developed in the MATLAB suite. The *sortie_subplot_stats* algorithm then calculated a density, biomass (Mg/ha) and an average canopy height for each subplot.

4.3.3 Probability Density Functions

In order to generate probability density functions of the SORTIE simulations to compare to the field data and validate the model, each SORTIE Prospect Hill Plot was run 100 times. The *sortie_assemble_forest* program assembled these 100 hectare plots into a virtual forest while still assigning each tree to its respective 25 x 25 meter subplot, with each plot containing 16 subplots. The resulting 1600 subplot measurements were then fit to probability density functions in order for statistical evaluation (p </=0.05). The large sample sizes allow for more robust conclusions of the evaluation techniques described in the model performance section.

4.3.4 Statistical Testing

To evaluate the accuracy of the modeled plots, two sample t-tests were used to verify no significant difference in the average biomass and density at the subplot level between the modeled plots and the field measurements. That is to say the field measurements could have been sampled from the probability density function of the simulated results. To be considered statistically similar, the test had to fail to reject the null hypothesis of no difference in means (p </=0.05).

The second phase of statistical evaluation involved passing each set of simulations though an F-statistic variance test to test for similar variances in biomass distribution and tree density at the subplot level in both the simulations and the field measurements. For the variances to be considered statistically similar, the test had to fail to reject the null hypothesis of no differences in variance (p </=0.05).

4.3.5 Sensitivity Analysis

In order to verify the correct functioning of the model, a sensitivity analysis was conducted on the growth and mortality parameters of the current SORTIE parameterization. The method used was a Monte Carlo simulation of 1000 runs of the model, randomly and independently varying both the growth and mortality parameters in each run. The Monte Carlo simulation was broken up between all ten Prospect Hill plots, with 100 simulations assigned to each. A Python script was used to generate two lists of random numbers, one for the growth parameter and one for the mortality parameter.

For the growth parameter, the mean growth rate of 0.1 cm/year was used, with a standard deviation of 0.02 cm/year. The random numbers were drawn using a Gaussian distribution. Of the resulting 1000 numbers, 95 per cent fell within the range of 0.06 to 0.14 cm/year, which we consider to be a realistic range for the growth parameter based on stand ages and current field measured average diameter (Hall, 2005, Cook et al 2011). Python uses the Mersenne Twister as its core random number generator. It produces 53-bit precision floats and has a period of 2**19937-1. The Mersenne Twister is one of the most extensively tested random number generators in existence (Matsumoto and Nishimura, 1998).

The set of random mortality values was generated using the same algorithm. The mean input into the random number generator was a 0.018 probability of mortality in any given year, with a standard deviation of 0.005. Of the resulting 1000 numbers, 95 per cent fell within the interval of 0.012 and 0.023, which we considered to be a reasonable range of mortality probability from based on the densities and average DBH of the field measured plots (Cook et al 2011).

The lists of both randomly generated growth and mortality values were combined, and a unique set of both parameters assigned a run number. The value sets were divided into subsets of 100 and assigned to one of the ten Prospect Hill plots. The growth and mortality values were entered into the SORTIE parameter set for each individual run, and the simulations were batched into sets of 100 runs. The resulting average biomass at the subplot level was extracted using the MATLAB SORTIE processing suite, and attributed to the unique growth and mortality parameter set. These values were then fed into the R statistical package, and a scatter plot produced for each Prospect Hill plot and all 1000 runs. Linear regression was used to fit a model to the Monte Carlo simulation of each of the Prospect Hill plots.

4.4 Objective: Visualizations

The final objective was to create a three dimensional (3-D) model of the forest using realistic canopy structures. For each tree in a simulated plot, the detailed output files generated by the SORTIE program included the species, x coordinate, y coordinate, DBH, tree height, canopy depth, and biomass. These data for each tree are stored in .xml format.

The detailed output .xml files from the final time step of one model run were fed into the suite of MATLAB programs to process the trees into an array. The *sortie_forest* program translated the tree height, diameter of the trunk, canopy depth and canopy width into a virtual three dimensional tree. To produce realistic canopy profiles for each tree species, profile views of each of the nine tree species were scanned from *Trees of the Northern United States and Canada* (Farrar 1991). The profile of the canopy was traced and translated into a vector. This vector was then rotated cylindrically to produce an idealized canopy structure for each tree species.

In order to validate the realism of the visualizations created in MATLAB, in 2012 the canopy profiles generated by the simulations were compared to new field measurements. The exact spatial locations of trees were mapped in plot PH3 subplot 1, along with the DBH, canopy height and depth. These measurements were inserted into the MATLAB visualization suite, for comparison to visualizations generated by SORTIE simulations.

5. RESULTS

Prospect Hill Plot 3 was the primary plot upon which this parameterization of SORTIE was trained. For clarity, the focus of this section will focus upon the results obtained from Prospect Hill Plot 4. Results from the remaining plots are included in Appendix 1.

5.1 History of Prospect Hill Plots and Model Initialization

The vector GIS layers were intersected with the DESDynI field plots, and the resulting polygons translated into the SORTIE grid system to initialize a model for each plot. The natural and anthropogenic histories both between and within these plots is very heterogeneous. Prospect Hill Plot 3 displayed the most homogenous pattern of disturbance, so it was chosen for initial estimation of modeling parameters.

According to historical records, Prospect Hill Plot 3 was abandoned in 1980 (Foster and Boose 1999; Figure 5). A small portion of the plot is listed in the agricultural abandonment map as being abandoned "pre-1908", which implies uncertainty in the historical records (Foster and Boose 1999). For modeling purposes homogenous abandonment in 1870 is assumed. The portion of plot PH3 which intersects the "pre-1908" polygon is negligible, and probably within the margin of spatial error of the map.

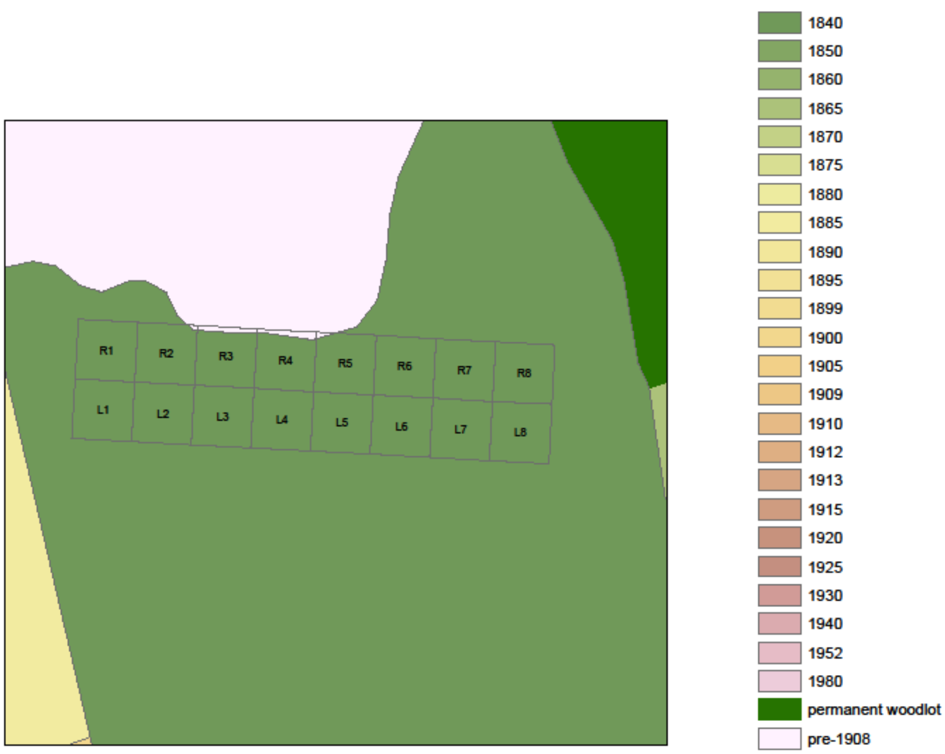


Figure 5. Prospect Hill Plot 3 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown

The 1938 hurricane damage map indicates damage of 21 -50 % in all of subplots 6, 7 and 10 through 16, and similar damage in the half of subplots 5, 6 and 9 (Hall 2005). This damage was translated as a partial harvest of 50% to the SORTIE grids representing these subplots (Figure 7). The higher end of the damage range reported within these subplots programmed into the model because all damaged timber was harvested in the period 1931 to 1941 (Foster 1992, Rowlands 1938).

Figure 6. Prospect Hill Plot 3 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.

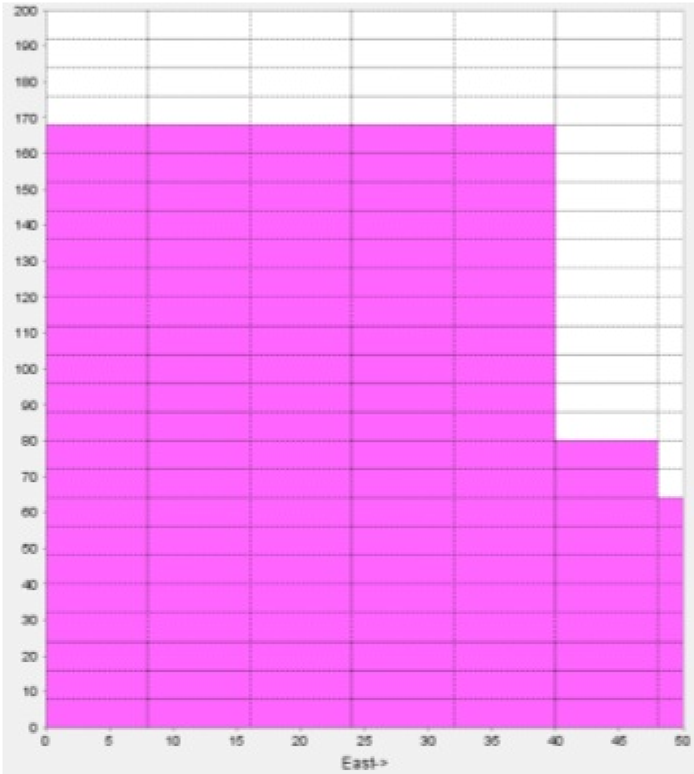


Figure 7. PH3 SORTIE harvest interface grid. A partial harvest was applied to the south eastern subplots. Distances in meters.

5.2 Modeling

The final model parameters were run for 100 simulations for each simulated field plot. Because SORTIE is a stochastic model, the results varied between each simulation. As statistical validations were only conducted on the final time step of each simulation it is important to make sure the dynamics of the simulation through time are capturing the desired historical events modeled within each plot. The effects of the 1938 hurricane and the thinning in 1949 are very apparent in the model of PH3 (Figure 8). Visual validation was used to determine if the simulations were capturing the relative dynamics of historical events. These results were also used to determine if the simulations were producing the correct relative dominance of tree species as measured in the field.

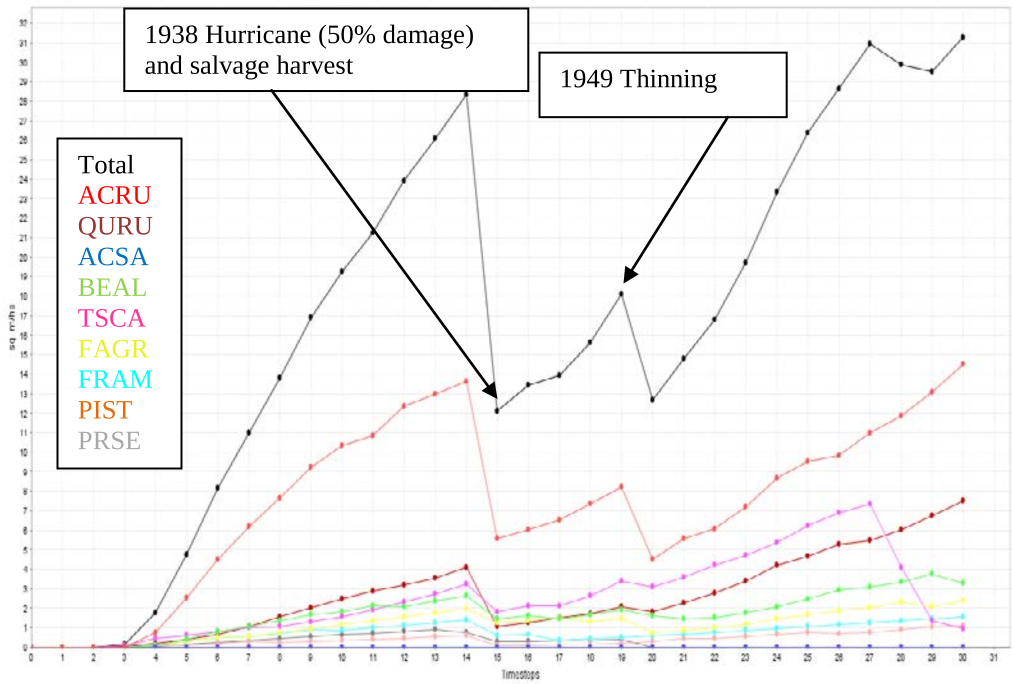


Figure 8. Modeling run of Prospect Hill PH3. X axis indicates time step, with T_0 corresponding to 1850. Y axis indicates basal area (m^2).

5.3 Evaluation of Model Performance

<u>5.3.1Plots outside Prospect Hill</u>

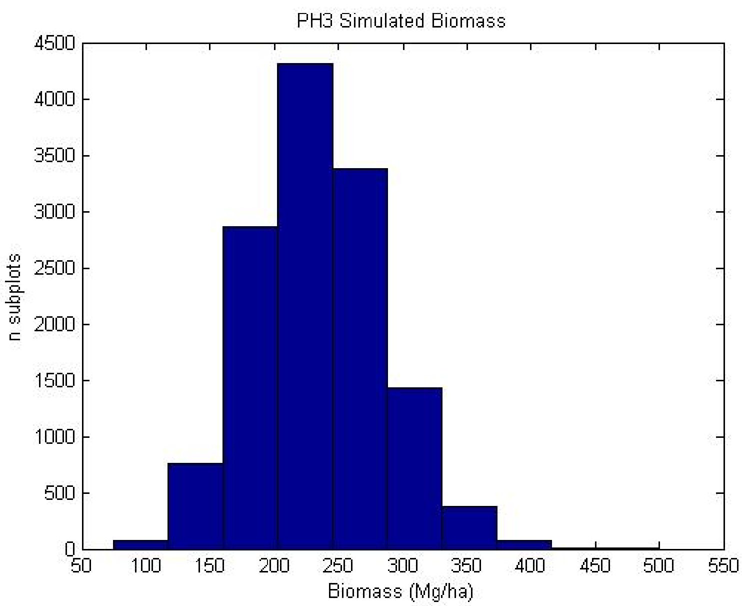
For Tom Swamp Plot 1, which was most like Prospect Hill Plot 2 , T_0 was chosen to be 1890, meaning the simulations were run for 24 time steps. The hurricane damage map does include Tom Swamp, and the resulting damage was 50 per cent uniformly to the entire plot (Hall 2005). Tom Swamp is also included in the silviculture treatments map, so the thinning in time step 14 was included in the model (Hall 2005).

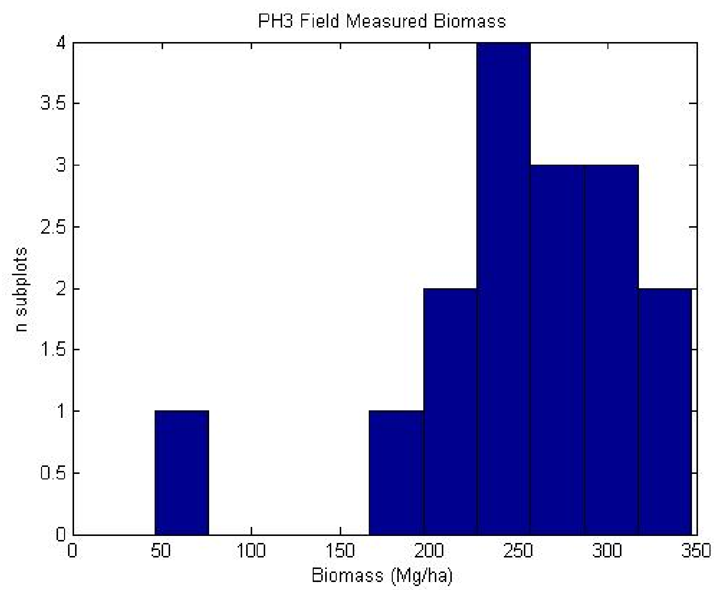
Tom Swamp Plot 2 most resembles Prospect Hill Plot 3, therefore 1850 was chosen as the initial time step, and the model run for 30 time steps. The hurricane damage within the plot was also uniformly 50 per cent. There were two thinning treatments in portions of the plot in time steps 20 and 25 (Hall 2005).

Slab City Plot 1 seems to match up with Prospect Hill Plot 5, and 1870 was chosen for T_0. Slab City also falls within the coverage of the hurricane damage and silviculture treatments layers. The hurricane damage to the plot was uniformly 90 percent and there was a thinning in time step 22 (Hall 2005).

5.3.2 Probability Density Functions

Prospect Hill Plot 3, which was the first plot to be parameterized and validated, is included here to illustrate the process of validation. The distribution appears to be Gaussian, according to the MATLAB function FITDIST. This indicates the field measured distribution could have been sampled from the simulated distribution.





Figures 9-10. Probability Density Function of Prospect Hill Plot 3 simulations. A fitdist test shows the field measured distribution could have been sampled from the Gaussian distribution formed by the simulations (p</=0.05). Number of simulated subplots = 1600. Number of field measured subplots = 16.

5.3.3 Statistical Testing

The results of the statistical evaluation show that most of the field measurements could have been drawn from the probability density function formed by the combined simulated runs for each plot (Table 5). All simulations produce the realistic biomass estimates. Within the Prospect Hill Plots, PH1, PH2, PH3 and PH8 display the best fits. Of the evaluation plots, SC1 displays the best fit. TS1 and TS2 produce the correct biomasses, however under predict the densities. The variance test seems to fail the most often, however when assembling the simulations and calculating their statistics, all the runs are averaged, which tends to lower the variance. If a single run is selected from the 100 total runs, the variance test fails to reject much less often.

Plot	Field Biomass	Simulated Biomass	Biomass p	Field Density	Simulated Density	Density P	Variance P	Simulation runs
PH1	304.8133	330.1012	0.2130*	937	978.69	0.6253*	0.1103*	100
PH2	256.0364	239.9719	0.2887*	967	851.30	0.6992*	0.5462*	100
PH3	248.7106	237.4050	0.4104*	476	460.76	0.5233*	0.0773*	100
PH4	271.6172	265.7428	0.5979*	551	554.27	0.9187*	0.00097	100
PH5	214.401	257.5401	0.2151*	702	560.31	0.00820	0.4952*	100
PH6	110.2564	135.4696	0.0941*	817	797.73	0.8258*	0.01860	100
PH7	279.5748	284.5996	0.7183*	834	779.78	0.2335*	0.02800	100
PH8	218.4617	203.6723	0.3866*	489	439.79	0.0669*	0.3726*	100
PH9	189.7594	192.2628	0.7518*	574	602.18	0.2497*	1.66E-23	100
PH10	134.5962	129.9585	0.8072*	785	664.92	0.1739*	0.03200	100
TS1	230.2684	233.0315	0.8497*	874	626.76	2.1e-017	0.01460	100
TS2	231.1013	240.3166	0.4609*	520	459.72	0.00630	0.4166*	100
SC1	192.0549	213.9481	0.2294*	762	823.50	0.2812*	0.0616*	100

Table 5. Statistical Evaluation. A two sample t-test was used to compare the simulated and field biomass and density measurements at the subplot (25m x 25m) level. An F-statistic variance test was used to compare the field measured and simulated variances.
 * Significantly similar ($p>0.05$)

5.3.4 Sensitivity Analysis

The Monte Carlo simulations show the model to be functioning as intended, and the ranges of the growth and mortality parameters selected to be within the range of the model's functioning (Figures 25 and 26). All lines, with the exception of the biomass parameter for PH6 in green, display a slope of less than one. This means that for a 10 percent change in the parameter, there is a less than 10 percent change in the response variable. This is due to the negative feedbacks in the model. PH2, in red, seems to have a noisy response to changing the growth parameter, as the coefficient of determination is only 0.06, indicating dispersed residuals.

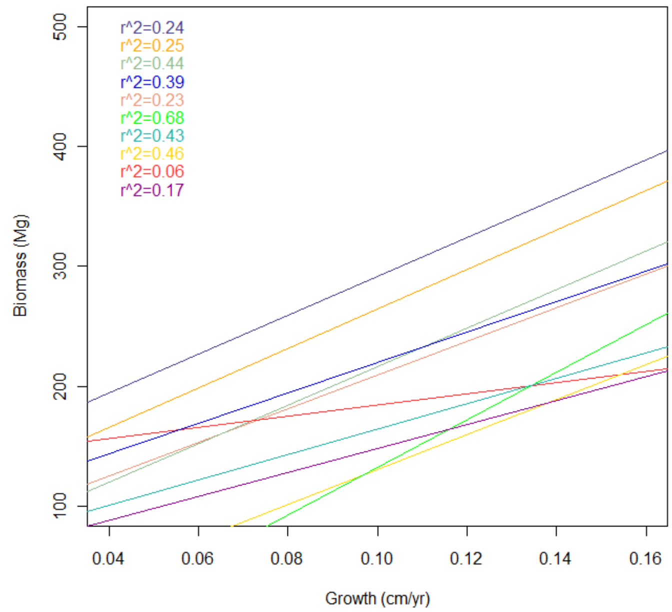


Figure 11. Sensitivity analysis of SORTIE growth parameter. Each line represents the best fit linear regression through the 100 simulated points for each Prospect Hill Plot. The r-squared coefficient of determination is included to indicate the fit of each line.

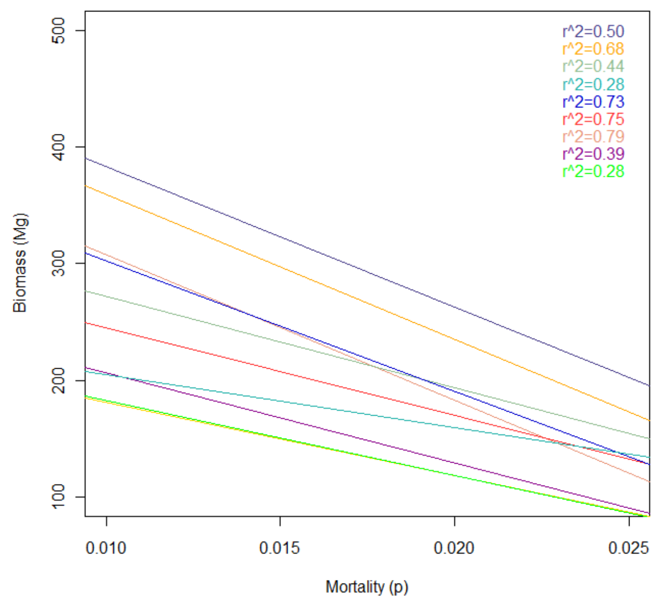


Figure 12. Sensitivity analysis of SORTIE mortality parameter. Each line represents the best fit linear regression through the 100 simulated points for each Prospect Hill Plot. The r-squared coefficient of determination is included to indicate the fit of each line.

5.4 Visualizations

Each plot measures 50 meters by 200 meters. Red canopies represent red maples, while brown canopies represent red oaks. Dark green canopies represent hemlocks while grey canopies represent white pine. Light green canopies represent yellow birch. The spatial configuration of the trees in the visualizations displays the influence of the historical natural and anthropogenic disturbances on current forest structure, biomass distribution and tree density.

The visualizations demonstrate the effects of differential disturbance patterns on the different plots. The northern section of Prospect Hill Plot 3 more resembles a late succession forest, with larger, less densely spaced trees Figure (12). These visualizations also demonstrate the high level of variability of both tree density and biomass distribution at the 25 x 25 meter scale.

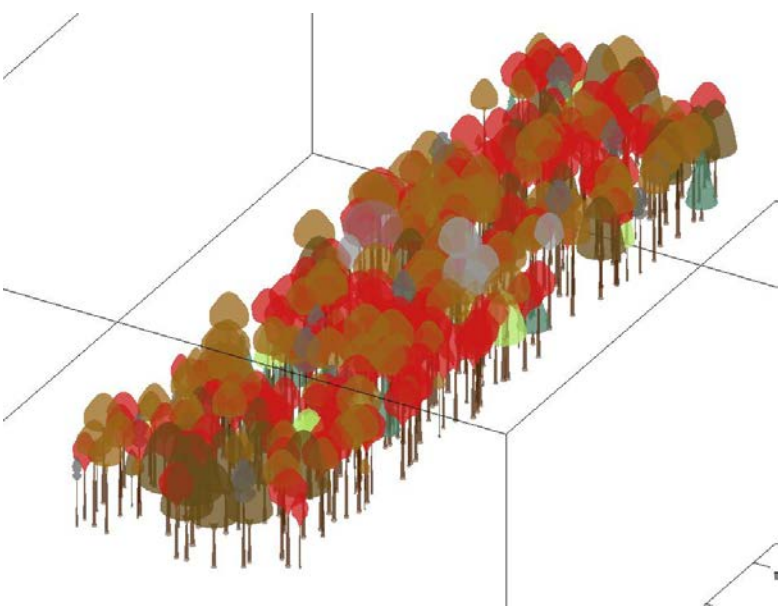


Figure 13. Visualization of Prospect Hill Plot 3 simulation.

The visualization of the 2012 field measured PH3 subplot 1 shows the simulationed visualizations are producing realistic canopy structures (Figure 14). The SORTIE model seems to assign canopy heights which are too deep. This indicates the canopy depth allometric parameters of SORTIE need to be adjusted for different sites.

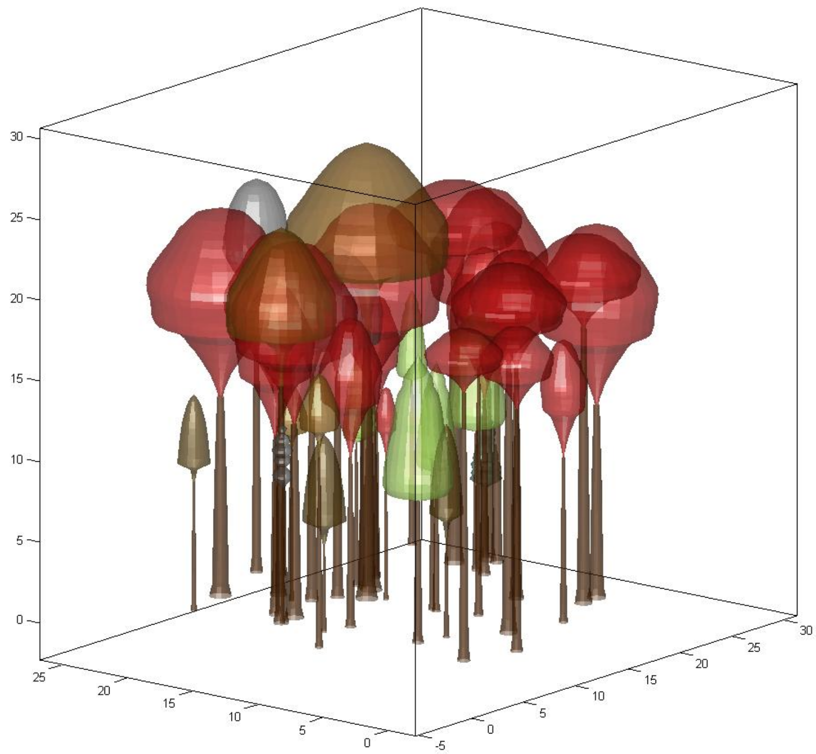


Figure 15. Visualization of 2012 field measurements for PH3 subplot 1.

6. DISCUSSION

Individual-Based Modeling offers a way to investigate the above ground biomass distribution within a forest stand based upon differential ecological and historical phenomena which had led to a specific distribution. This analysis will help to inform us of both how different parameters affect biomass distribution and the impacts of changing these parameters. This form of modeling will also help those attempting to develop techniques for remote sensing of biomass and carbon in addressing sources of error and uncertainty in biomass estimation.

We have parameterized and validated the SORITE model of forest succession to reflect the how differential ecological and historical parameters have led to the specific biomass distribution in ten Prospect Hill plots. This parameterization tells the history of each plot. To further this investigation, we have used the output from this modeling to produce three dimensional representations of the forest structure within these plots.

In lieu of growth records for individual trees, or estimating rates based upon measuring the width of rings obtained from core samples, we have developed an approach to solve this parameter value by combining historical records with current field measurements. If it is known when a plot started to grow, all the disturbances which happened within the plot, and the current biomass and tree density, we can estimate the growth parameter with a degree of confidence.

We have also developed a method to account for the lack of soil and moisture mechanisms in the SORITE model (Pacala 1996) and model the relative competiveness of each tree species. By examining the current stand composition, examining historical land use maps (Foster 1992), and researching the relative competitiveness of tree species on these land use types (Motzkin 1990), we were able to alter the asymptote and slope of the growth response curves for each species to produce realistic results. Through an iterative process we were able to produce a set of relative slope response parameter values which produce the correct species compositions in all plots, even the three upon which the model was not trained.

The sensitivity analysis using Monte Carlo methods to test the effects of randomly perturbing the growth and mortality parameters shows the model to be functioning properly. Increasing the intrinsic growth rate parameter (cm/year) by ten percent results in an increase of plot biomass by less than ten percent in all Prospect Hill plots except for PH6 (Figure 11). The reason for the odd functioning of PH6 will be discussed later. Increasing the intrinsic probability of the mortality parameter in the simulations by ten percent in all plots results in a decrease of less than ten percent biomass (Figure 12). These results indicate proper negative feedback loops are functioning to dampen the changes in these parameters. The regressions of the points for each plot show a linear relationship between the independent and response variables.

The results of the Monte Carlo simulations give realistic ranges for the growth and mortality parameters in future SORTIE simulations. However the factors controlling tree mortality operate at a much larger scale then those controlling the growth parameters. These smaller scale variables such as elevation and soil type are not included in the SORTIE modeling framework (Pacala 1996). This leads to the argument of assigning the average mortality rate derived from the Monte Carlo simulations to all the plots, while varying the growth conditions about a much larger prior probability density function to account for factors not included in SORTIE.

While model performance evaluation shows most simulations seem to match the field data, not all plots are fit perfectly (Table 5). The biomass from the field all fall within the prior distribution created by the simulations. However, some of the densities and variances are not fit perfectly. The danger in over fitting a model lies in how useful it will be outside the specific area in which it was parameterized. A model which has been over fit will likely impose an understanding of a system which in reality only represents the data upon which it was trained. This study makes every attempt to avoid this pitfall; however the detailed nature of the data used to train this model may in fact lead to a site specific model.

A number of available data were used to parameterize this model. From land use history, detailed silviculture records, hurricane damage maps and extensive field observations, the rich historical and ecological record that exists for the Prospect Hill plot of Harvard Forest allows for a fine level of historical resolution in modeling. Though this process, however, we found clear path dependence in forest makeup, density and biomass distribution from historical events.

The increase in models that represent the functioning of complex adaptive systems has led to an increased awareness of path dependency and multiple equilibria in ecological systems (Pahl-Wostl 1995). This path dependence stems from negative and positive feedbacks. For example, just changing the intensity of one simulated thinning of a plot by ten percent can lead to more than ten percent change in tree density and biomass in the final time step of plot PH3(Figure 8). Chaotic systems are characterized by sensitivity to small perturbations, implying that small disturbances have disproportionally large and long lived effects (Phillips 2004). Similar path dependence also applies to the simulated hurricane damage within Prospect Hill Plot 3 (Figure 8).

Path dependence related to historical events both natural and anthropogenic highlights a source of uncertainty and error in both forest modeling and biomass estimation that is relatively unexplored. Prospect Hill plot 6 was the last plot to be parameterized and validated, due to the high density and small average diameter of trees within the plot (Cook 2010). Within the data used to train the other nine plots there were no historical factors or events to explain the forest composition of PH6. However discovery of a natural disturbance map indicated a forest fire affected the northeastern section of Prospect Hill in 1957 (Foster and Boose 1999). Of the ten research plots the fire only affected PH6; in fact PH6 was the only field plot affected by any recorded natural disturbance (the hurricane damage within Harvard Forest is recorded separately). Without simulating the fire, there is no way to produce results similar to the field validation measurements of PH6 without altering many other parameters in the SORTIE model. As red oak is more tolerant to fire than the other species present, and more quickly to regenerate after (Farrar, 1991), it has come to dominate the plot (Figure 16).

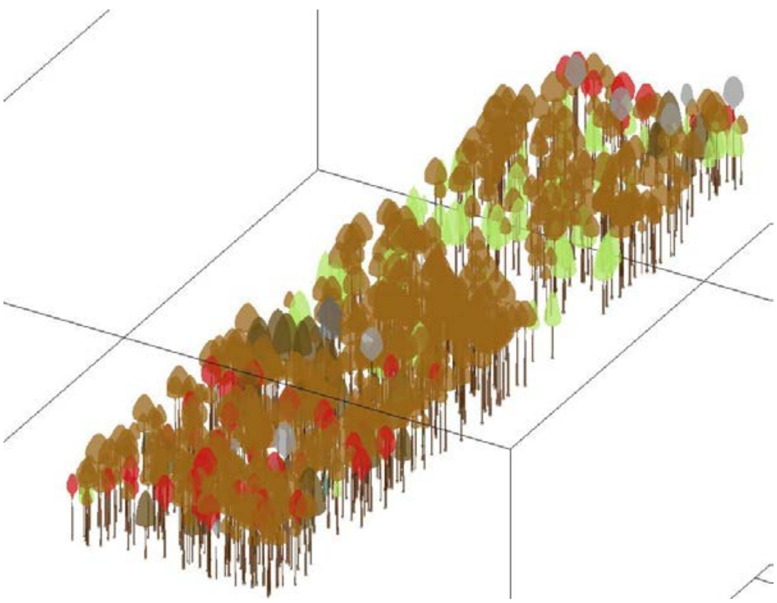


Figure 16. Visualization of Prospect Hill Plot 6.

The path dependence inferred by historical anthropogenic events is best illustrated by the three dimensional visualizations. The visualization of the PH5 simulation shows the effects of differential abandonment, the path dependence caused of different starting points. It is apparent the portion of the plot abandoned in 1870 is more dominated by late successional species such as white pine, whereas the portions of the plot abandoned later are dominated by red oak (Appendix Figure 26). The effects of a thinning are most evident in the visualization for PH2, in which a thinning took place in 1957 in the north portion of the plot. The trees in the north part of the plot appear much less densely configured than the southern portion, which is also dominated by more red oaks while the rest of the plot is hemlock dominated (Appendix Figure 28). This path dependence poses a problem in attempting to model forests, especially those with a less complete history than the Prospect Hill Tract.

This limitation is evident when the SORITE parameterization of this is applied to plots outside Prospect Hill. Although the simulations of the Tom Swamp plots and Slab City plots accurately model the biomass measured in the field, the model systematically under-predicts the density (Table 5). This is most likely due to the fact that although the hurricane damage and silviculture treatments within these plots are known, the starting time point and other natural disturbance factors are unknown. These simulations may be missing some factor which causes density of the trees within the plots to increase and lies outside the historical record. In addition, it is well-known that diverse combinations of height, BA and density can produce similar biomass quantities. Thus biomass may be a more realistic estimate for similar modeling estimates, or for remote sensing-based estimates, than density alone.

Exploration of path dependence in biomass distribution and forest density brought about by a specific, spatially explicit history of natural and anthropogenic disturbances is now possible through advances in both computing power and three dimensional spatially explicit forest modeling. Obrien et al have investigated the effects of differential hurricane regimes on tropical forests, although they used the ZELIG model with hypothetical forests and differential hypothetical hurricane regimes (Obrien et al 1992). Desai et al parameterized the ED model with ecological, forest inventory, and historical land use observations in an intensively managed forested landscape in the upper Midwest United States to study the effects of disturbance on carbon cycling (Desai et al 2007). However, as the ED model only considers canopy height (Moorcroft et al 2001), the modeling methodology of this paper allows for more detailed research into the effects of disturbance on three dimensional canopy structures produced by differential disturbances.

Most spatially explicit studies of path dependence in complex adaptive systems are limited to describing the configuration of city growth in urban planning (Atkinson and Oelson 1996, Wilson 2000), landform evolution (Perron and Fagherazzi, 2011) and the role of initial conditions and divergences from those in ecological systems (Phillips, 2004).

Motzkin et al studied the impacts of initial conditions, natural and anthropogenic disturbances on the specific species compositions of plots within Prospect Hill (1999). They were able to correlate the dominance of specific tree species to different disturbances. These findings were used in our parameterization to make red maple and red oak more competitive. Uriate et al (2007) used the SORTIE model to investigate the effects of increased hurricane frequency on carbon storage over the entire southern New

England region. They concluded over the short term increased hurricane damage will lead to a drop in biomass and carbon storage, however the long term effects depend on the fate of the downed timber; whether it is salvaged or not. However this study only projected the current SORTIE model into the future and extrapolated the potential effects of increasing hurricane frequency, with no validation to back up their findings. This highlights the need for data driven modeling research into the dynamics of path dependency and multiple equilibria in above ground biomass distribution and tree density brought about by natural and anthropogenic disturbances in forest ecosystems supported by historical records and field measurements.

The Harvard Forest (USA) Nitrogen saturation experiment reports increased aboveground production in hardwood stands with elevated nitrogen levels, while red pine plantations under elevated nitrogen displayed a decrease over the period (Magill et al 2004). These show the importance of soil nitrogen in determining above ground biomass production. The version of SORTIE used in this study (Pacala 1996) does not consider soil nitrogen, although a grid 'labeled secondary resource' could be used to account for the element. However, the field data was collected for remote sensing purposes and soil samples were not taken (Cook 2010). Including nitrogen cycling in this model would help to produce its accuracy.

The depth of the historical record for Prospect Hill belies both strength and weakness to the model parameterization of this study. The detail with which the model was parameterized will allow for confident forecasts of forest dynamics into the future for Prospect Hill under different scenarios. This may prove useful in modeling the effects of the *Adelges tsugae* (Hemlock Woolly adelgid), an invasive pest which causes high mortality in hemlocks (Orwig & Foster, 1998). As hemlock is one of the most abundant, long lived and shade tolerant species in Northeastern forests (Rogers, 1978), the unique role it plays in these ecosystems will cause its loss to dramatically alter the landscape of this region.

Another possible future scenario this model may prove useful in forecasting is the effects of increased atmospheric carbon dioxide levels on biomass distribution and carbon storage within Prospect Hill. The path dependence caused by disturbance events encountered in this study may prove useful in generating management plans to maximize the amount of carbon stored in these forest stands. The actual modeling of actual hurricane damage and it measured results in 2009 will allow for more confident predictions about the effects of increased hurricane frequencies. However, the site specificity of the data with which this model was trained makes any forecasting only applicable to the Prospect Hill site, and the specific forest dynamics within. The general model can be applied outside the plot, but the actual forest structure generated by the model may not reflect the dynamics of that specific forest.

Due to the high level of spatial resolution included in this model however, its results may prove useful in developing an approach to remotely sense biomass using lidar. In forests, the physical propagation, extinction, and scattering of light is in part a function of the depth, width and structure of forest tree crowns and gaps between them. Spatially explicit 3D simulated forests from IBMs can be coupled with a lidar scattering model.

First the three dimensional scenes generated during the visualization step of this study would be divided into cubic cells of a size corresponding to the resolution of the lidar system. This process is known as voxelization. The resulting cells are attributed to the different structural components of the forest and assigned different densities. A physics-based scattering algorithm is used to estimate the lidar backscatter from each of these cells (e.g. Sun and Ranson, 2000), as the energy from the lidar pulse is either reflected from the cell or transmitted through the cell to those below.

Although Sun and Ranson successfully modeled the 3D lidar returns over a simulated Maine forest in 2000, they did not predict forest biomass. The IBM model used in the study (ZELIG) also did not provide spatially explicit tree locations or crown metrics, and the shapes assigned to tree canopies were simple cones and cylinders. Hutt et al (2010) used lidar measured canopy heights to initialize the Ecosystem Demographic (ED) model developed by Moorcroft (2001) to predict biomass. However, the study sites used for these studies had a great deal of heterogeneity of canopy heights, and the ED model only tracks the height dimension of canopies. Within certain forests, stands of the same height have been measured to have differing biomasses. Previous studies of remote sensing biomass through lidar have reported high correlation coefficients of determinations (~R^2=0.8), however have included a wide range of forest heights (Nelson et al 2003). The approach we have developed to model the Prospect Hill tract was designed specifically to address the shortcomings of these previous studies. The results from our SORTIE modeling have a much higher degree of spatial resolution than either the ZELIG or the ED models. The shapes used to give the crowns a three dimensional shape in this study are also much more realistic than the simple cones and cylinders used by Sun and Ranson.

In order to better support this mission, a more generalizable form of this model needs to be developed. Most forests lack historical records of disturbances, whether natural or anthropogenic. This proves very troublesome for accurate modeling of forest structure and dynamics due to the path dependence brought about by these disturbances. It seems some form of average disturbance needs to be applied to a given plot where the historical disturbance regime is unknown and/or poorly documented.

Another aspect of the forest structure which is important to the remote sensing of biomass is the density of trees within the forest. While the end goal is to obtain the biomass, the density of the forest is also very important. Two forests which have similar biomasses can display greatly differing tree densities (Table 5). Further research is needed to determine if the structure of a lidar return can be used to differentiate forests of different densities. This would help the modeling process, as the path dependence inferred by disturbance seems to be more directly related to density than biomass per se. A thinning more recently in time seems to decrease density while a thinning farther in the past seems to increase density. If lidar could differentiate between differing densities, the model could be more accurately tuned to reflect the actual forest structure and more accurate biomass estimate could be obtained.

The next step in creating an even more realistic model of forest structure is to use fractal geometry to generate virtual trees with realistic branching structures. The use of fractal geometry to produce realistic plant structures was pioneered by Arstid Lindenmeyer (1968). He developed a formal language, or grammar, known as L-systems which modeled the branching habits of plants. This method works by rewriting, for example *a* ->*ab* means *a* is replaced by sting *ab* and *b* ->*a* means *b* is replaced by *a*. This can be used to model the branching habits of trees, as trees demonstrate self-similarity as they branch. An angle of divergence must be specified to capture the angle at which a particular species branches (Prusinkiewicz and Lindermayer, 1990).

As an exploration, we have implemented an L-system in MATLAB to generate a sample fractal red maple (Figures 40 and 41). By measuring branching angles of red maples in the field plots, we determined the average angle of divergence to be 37.5 degrees. The implementation used the stochastic form of the L-systems, and branching angle and shoot length varied randomly about a mean based on a Gaussian distribution.

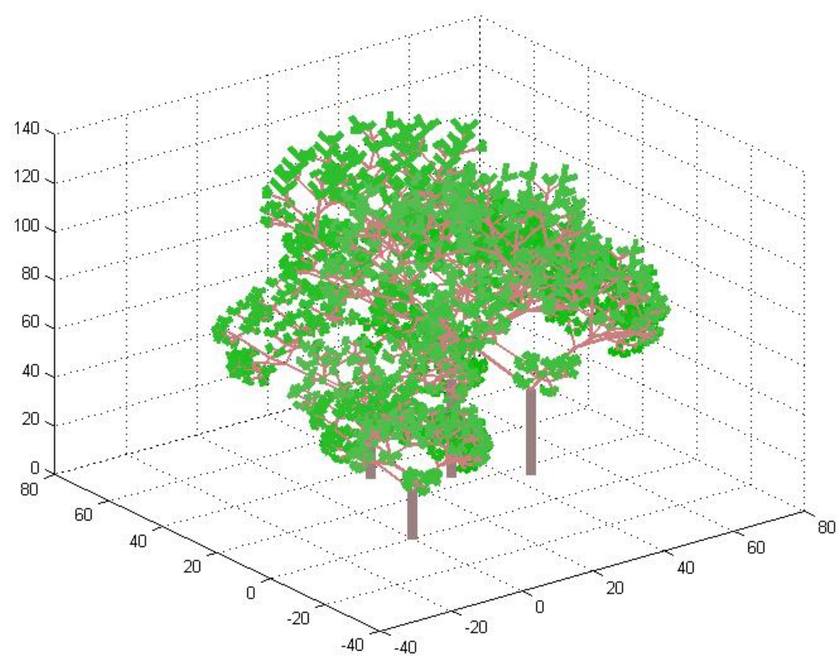


Figure 16. Simple fractal red maples.

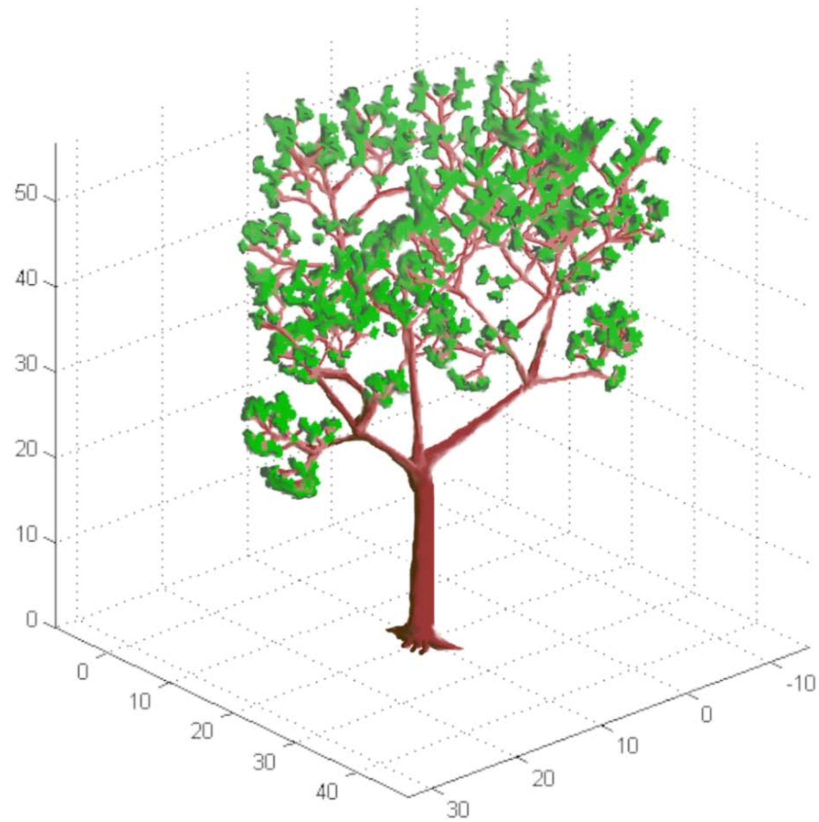


Figure 17. Sample fractal red maple with taper factor and shading.

Ideally, and with enough computational power, each tree output by the SORTIE simulations would be used to generate a unique fractal tree. Then the above ground biomass could be split into its different components (trunk, branches and leaves) and allocated accordingly. When put through the voxelation process, the result should provide a much more realistic model of forest structure. Hopefully this would greatly improve the accuracy of the simulated lidar return over this forest plot.

7. CONCLUSIONS

We developed a method to modify the SORITE model (Pacala 1996) to accurately estimate growth rate and relative competition between tree species by combining in absence of measurements of these parameters. This was made possible by combining detailed historical records of land use and disturbance with current field data. This method of parameter estimation is most useful when spatially explicit detailed historical data is available.

This technique has allowed us to construct one of the first spatially explicit historical models of biomass distribution within an actual forest. The visualizations can be used to create animations depicting the forest dynamics within these plots since agricultural abandonment. These plot histories will add to the already extensive records contained within the Harvard Forest Data Archive.

The accuracy of the simulations was statistically evaluated and a sensitivity analysis used to ensure proper functioning of the model and its assumptions. Through this modeling and the process of three dimensional visualization we discovered a high degree of path dependence in biomass distribution and tree density related to historical disturbance factors both natural and anthropogenic.

The nature of this path dependence will be of importance for future management decisions in the face of climate change and invasive pests. The level of historical detail upon which this model was trained will allow for more confident projection of future dynamics (Pacala 2010). Although the projections may be site specific to Prospect Hill, the results of recorded disturbances should allow for accurate modeling of these dynamics in other sites.

The results of this study also display the limitations in developing accurate forest succession models in stands which lack historical records of natural and human induced disturbances. This shows the need to development of the ability to develop mechanisms which simulate these disturbances to see if they produce accurate representations of current forest structures. The methods of constructing disturbance regimes employed in this study are an important first step toward this goal.

The three dimensional results of this modeling can be used to develop an approach for the remote sensing of biomass and carbon storage by lidar. This approach can be trained upon the Prospect Hill modeling results and hopefully prove useful outside the tract. Future areas of research include investigating the nature of the path dependency inferred by natural and anthropogenic disturbance and more realistic methods of creating three dimensional forest structures.

8. REFERENCES

Atkinson, G. & Oleson, T., 1996. Urban sprawl as a path dependent process. *Journal of Economic Issues*, 30: 609-615

Brown, S. 1999. Guidelines for Inventorying and Monitoring Carbon Offsets in Forest-Based Projects. Arlington (VA): Winrock International.

Botkin, D. B., Janak, J. F., and Wallis. J. R.: 1972, Some Ecological Consequences of a Computer Model of Forest Growth, *J. Ecol:*60, 849–872.

Bushing, R.T. 1991. Disturbance and the population dynamics of *Liriodendron tulipifera*: simulations with a spatial model of forest succession. *J. Ecol.* 83: 45-53.

Bursing, R.T., and Mailly D., 2004. Advances in spatial, individual-based modeling of forest dynamics. *Journal of Vegetation Science* 15: 831-842

Cook B., R. Dubayah, F. Hall, R. Nelson, J. Ranson, A. Strahler, P. Siqueira, M. Simard, & P. Griffith. 2011. NACP New England and Sierra National Forests Biophysical Measurements: 2008-2010. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. http://dx.doi.org/10.3334/ORNLDAAC/1046

Desai, A.R., Moorcroft, P.R., Bolstad, P.V., Davis, K.J. 2007. Regional carbon fluxes from an observationally constrained dynamic ecosystem model: Impacts of disturbance, CO2 fertilization, and heterogeneous land cover. *Journal of Geophysical Research-Biosciences* 112: Article number G01017.

Ek, A.R., & Monserud, R.A. 1974. Trials with program FOREST: Growth and reproduction simulation for mixed species even- or uneven-aged forest stands. In Fries, J. (ed) *Growth models for tree and stand simulation*, pp 56-73. Department of Forest Yield Research. Note 30, Royal College of Forestry, Stockholm, SE.

Farrar, J.L. 1991. *Trees of the Northern United States and Canada.* Fitzhenry & Whiteside, Markham, Ontario.

Fisher, R. T. 1921.The management of the Harvard Forest. Harvard Forest Bulletin1: 1–55

Fearnside, P. M. 1997. Monitoring needs to transform Amazonian forest maintenance into a global warming mitigation option. *Mitigation and Adaptation Strategies for Global Change* 2:285–302.

Fearnside, P. M. 2000. Uncertainty in land-use change and forestry sector mitigation options for global warming: Plantation silviculture versus avoided deforestation. *Biomass and Bioenergy* 18:457–468.

Foster, D. R. 1992. Land-use history (1730-1990) and vegetation dynamics in central New England, USA. Journal of Ecology 80: 753-772.

Foster, D.R. & Boose, E. 1999. Historical GIS data for Prospect Hill Tract at Harvard Forest 1733-1986. Harvard Forest Data Archive: HF055.

Hall B. 2005. Historical GIS Data for Harvard Forest Properties from 1908 to Present.. Harvard Forest Data Archive: HF110.

Hurtt, GC, Dubayah, R, Drake, J. Moorcroft PR, Pacala, SW, Blair, JB, Fearson MG. 2004. Beyond potential vegetation: combining lidar data and a height-structured model for carbon studies. Ecological Applications 13 873-883.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. 2003. National-scale biomass estimators for united states tree species. *Forest Science* 49 12–35

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. 2004. Comprehensive database of diameter-based biomass regressions for north american tree species. US Forest Service Northeastern experiment station research paper NE-319

Lindenmayer A (1968). Mathematical models for cellular interaction in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology* 18:280-289

Magill, A.H., Aber, J.D., Currie, W.S., Nadelhoffer, N.J., Martin M.E., McDowell, W.H.,Mellillo, J.M, and Steudler, P. 2004. Ecosystem response to 15 years of cronic nitrogen addidtions at the Harvard Forest LTER, Massachusetts, USA. *Forest Ecology and Management* 196: 7-28.

Matsumoto, M.; Nishimura, T. (1998). "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator". *ACM Transactions on Modeling and Computer Simulation* **8** (1): 3–30.

Mailly, D., Kimmins, J.P. & Bushing, R.T. 2000. Disturbance and succession in a coniferous forest of northwestern North America: simulations with DRYADES, a spatial gap model. *Ecol. Model.* 127:183 -205

Motzkin, G., Wilson, P., Foster, D.R., & Allen, A. 1999. Vegetation patterns in heterogeneous landscapes: The importance of history and environment. *Journal of Vegetation Science* 10: 903-920.

Moorcoft, PR, Hurtt, GC, & Pacala, SW. 2001. A method for scaling vegetation dynamics: The ecosystem demography model (ED). Ecological Monographs, 71 pp. 557-585.

Nelson et al., 2003 R. Nelson, M.A. Valenti, A. Short & C. Keller, 2003.A multiple resource inventory of Delaware using airborne laser data. Bioscience, 53 pp. 981–992.

Nilsson, M 1996. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment* 56: 1-7

Obrien, S.T., Hayden, B.P., and Shugart, H.H. 1992. Global climatic-change, hurricanes and a tropical forest. *Climate Change* 22: 175-190.

Orwig, D.A. & Foster, D.R. (1998) Forest response to the introduced hemlock woolly adelgid in southern New England, USA. *Journal of the Torrey Botanical Society* 125, 60-73

Pacala, S.W. 2010. Personal conversation. University of Michigan Environmental Spatial Analysis Laboratory, February 28, 2010.

Pacala, S.W., Canham, C.D., Saponara, J., Silander, J.A., Kobe, R.K. & Ribbens, E. 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecol. Monogr*. 66: 1-43.

Pacala, S.W., Canham, C.D., & Silander, J.A., Jr. 1993. Forest models defined by field measurements: I. The design of a northeastern forest simulator. *Can. J. For. Res.* 23: 1980-1988.

Pahl-Wostl, C., 1995. *The Dynamic Nature of Ecosystems: Chaos and Order Entwined* (Chichester: John Wiley).

Perron, J.T, and Fagherazzi, S. 2011. The legacy of initial conditions in landscape evolution. Earth Surf. Process. Landorms 37, 52, 63.

Prusinkiewicz, P., Lindenmayer A. 1990. *The Algorithmic Beauty of Plants.* Springer-Verlag.

Phillips, JD., 2004. Divergence, Sensitivity, and Nonequilibrium in Ecosystems. Geographical Analysis, Volume 36, Number 4 pp. 369-383.

Rasche, H. 1958. Temperature differences in Harvard Forest and their significance. Harvard Forest Paper 4: 153.

Rogers, R.S. (1978) Forest dominated by hemlock (*Tsuga canadensis):* distribution as related to site and post settlement history. *Canadian Journal of Botany*, 56, 843-854.
Shugart, H.H., 1984. *A theory of forest dynamics.* Springer-Verlag, New York, NY US.

Rowlands, W. 1939. As tudy of forest damage caused by the 1938 hurricane in the vicinity of Petersham, Massachusetts. *Harvard Forest Archives:* HF 1939-05.

Sun, G., & Ranson, K.J. 2000. Modeling Lidar Returns from Forest Canopies. *IEEE Transactions of Geoscience and Remote Sensing* 38: 6 : 2617 - 2626.

Shugart, H.H., and West D.C. 1997. Development of an Appalachian deciduous forest succession model and its application to assessment of the impact of the chesnut blight. *J Environ. Manag. 5.* 161-179

Smith, T.M. & Urban, D.L. 1988. Scale and resolution of forest structural patterns. *Vegetation* 74:143-150.

Stehman, S., Wickham, J., Smith, J., & Yang, L. 2003.Thematic accuracy of the 1992 national land-cover data for the eastern United States: Statistical methodology and regional results. *Remote Sensing of Environment* 86500–516.

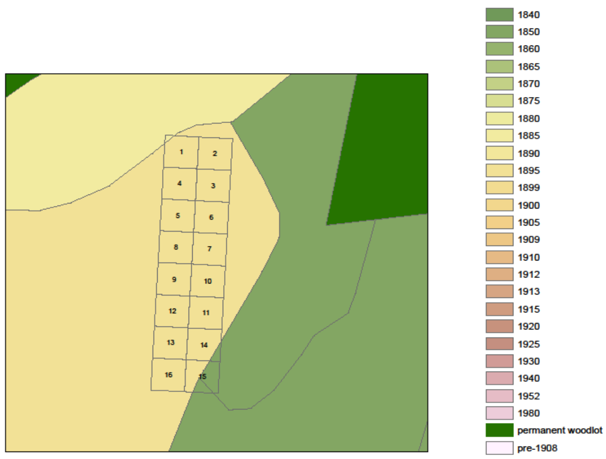
D.R. Zak, W.E. Holmes, K.S. Pregitzer. 2007. Atmospheric CO_2 and O-3 alter the flow of N-15 in developing forest ecosystems. *Ecology*: 88 (2007), pp. 2630–2639

Uriate, M. & Papaik, M. 2007. Hurricane impacts on dynamics, structure and carbon sequestration potential of forest ecosystems in Southern New England, USA. *Tellus* 59A: 519 – 528.

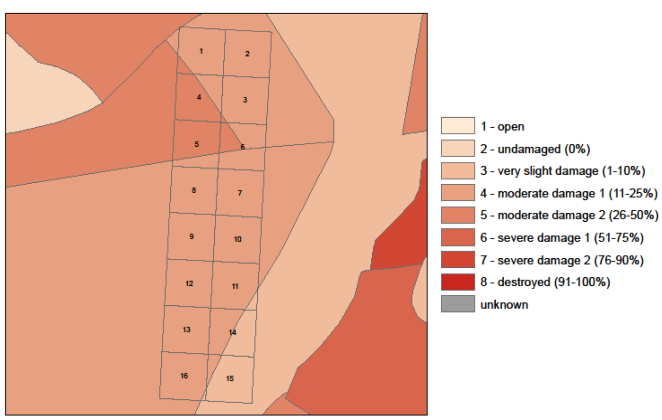
Urban, D.l., Bronan, G.B., Smith, T.M. & Shugart, H.H., 1991. Spatial application of gap models. *For. Ecol. Manage.* 42: 95 – 110.

Wilson, G.G., 2000. *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis* (New York: Pearson).

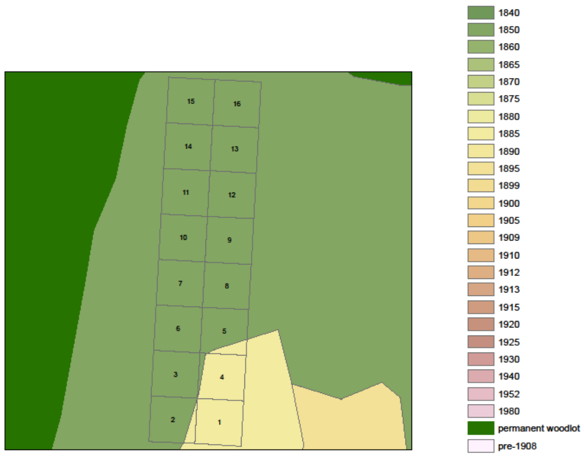
9. APPENDIX 1: ADDITIONAL FIGURES



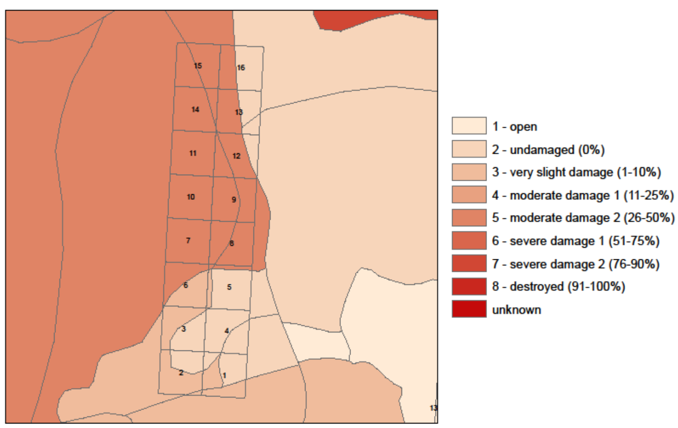
Appendix Figure 1. Prospect Hill Plot 1 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



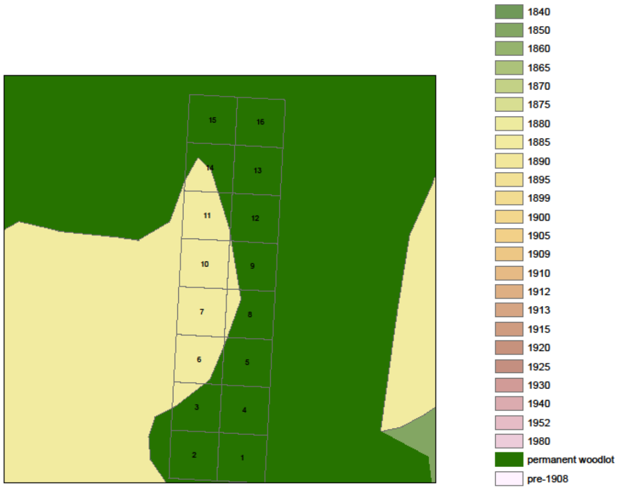
Appendix Figure 2. Prospect Hill Plot 1 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



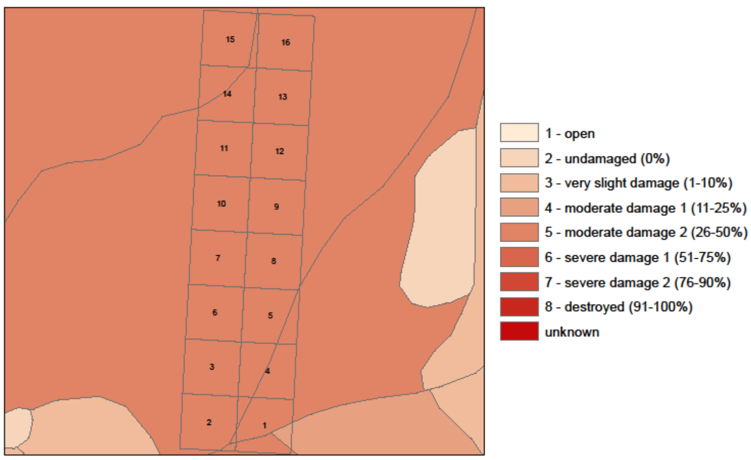
Appendix Figure 3. Prospect Hill Plot 2 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown



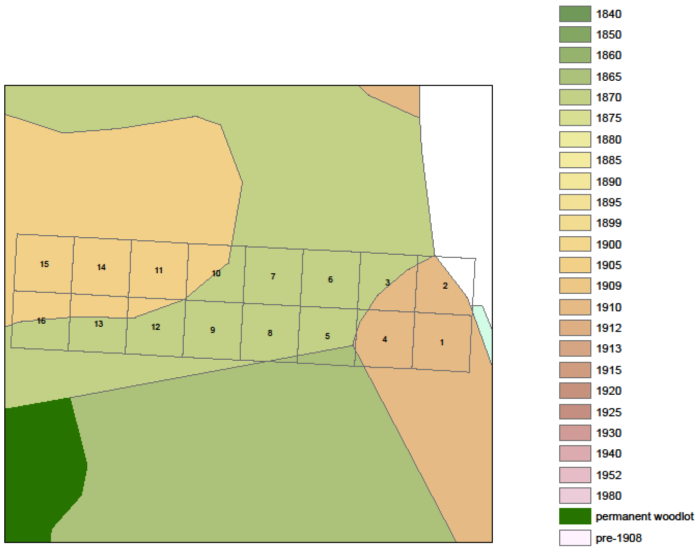
Appendix Figure 4. Prospect Hill Plot 2 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



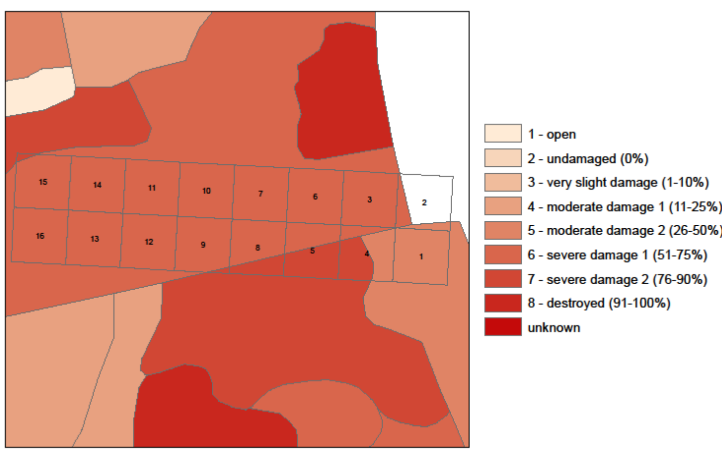
Appendix Figure 5. Prospect Hill Plot 4 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown



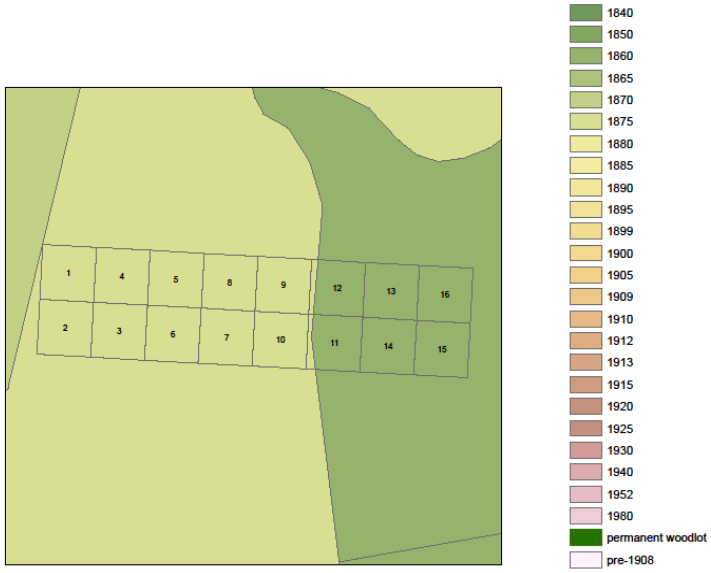
Appendix Figure 6. Prospect Hill Plot 4 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



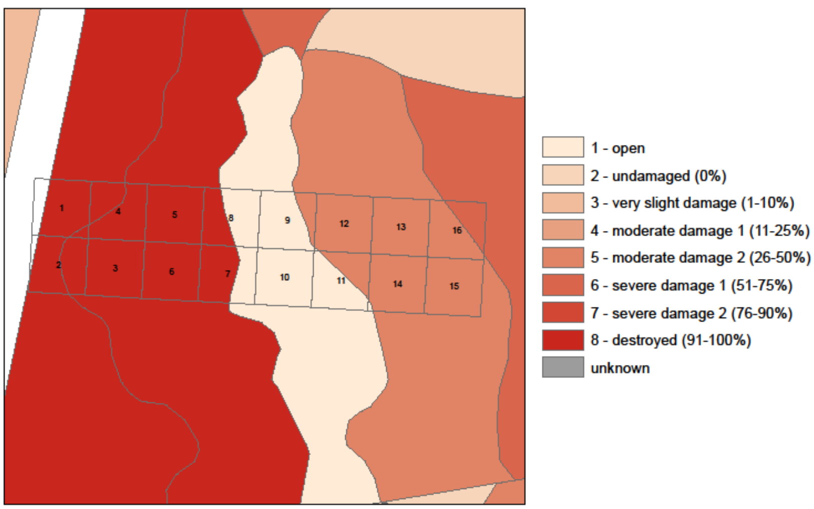
Appendix Figure 7. Prospect Hill Plot 5 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



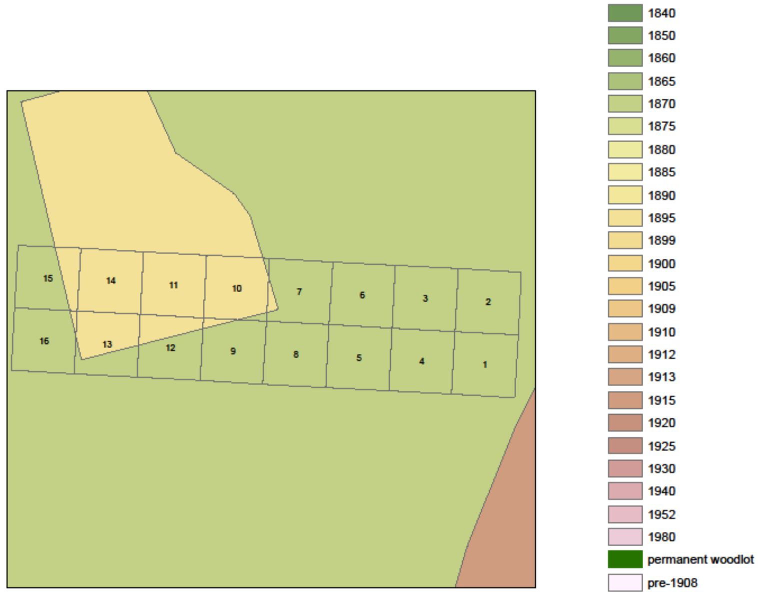
Appendix Figure 8. Prospect Hill Plot 5 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



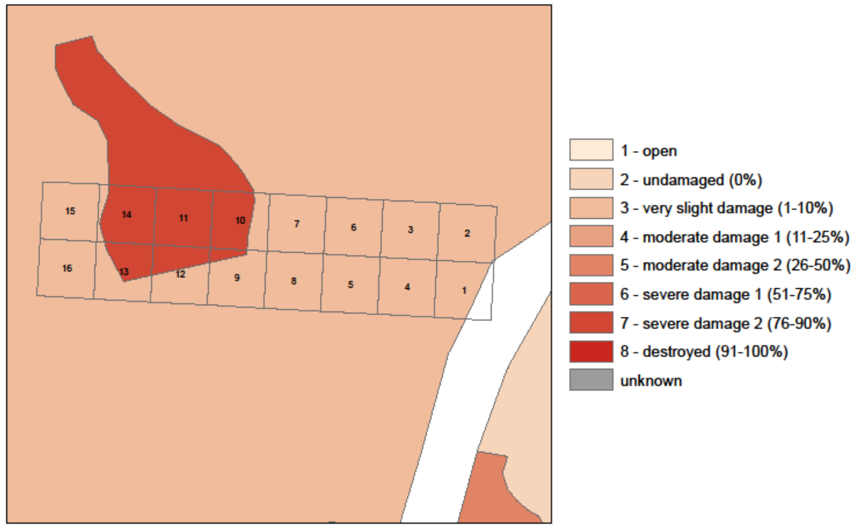
Appendix Figure 9. Prospect Hill Plot 6 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



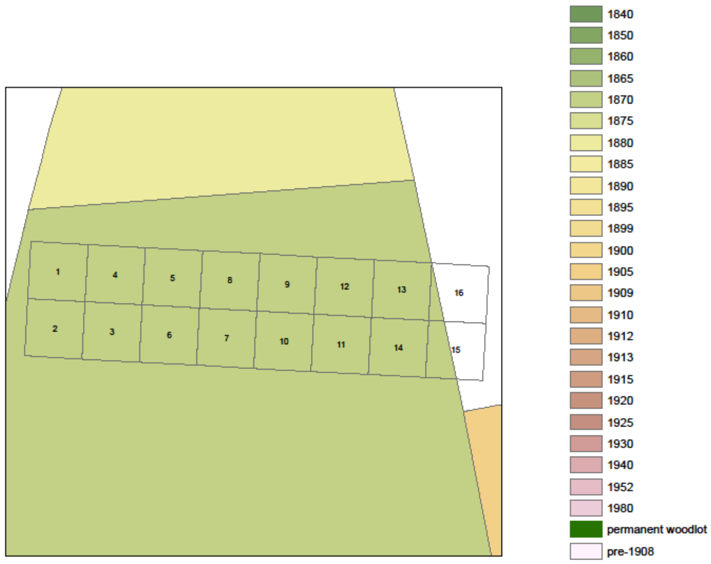
Appendix Figure 10. Prospect Hill Plot 6 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



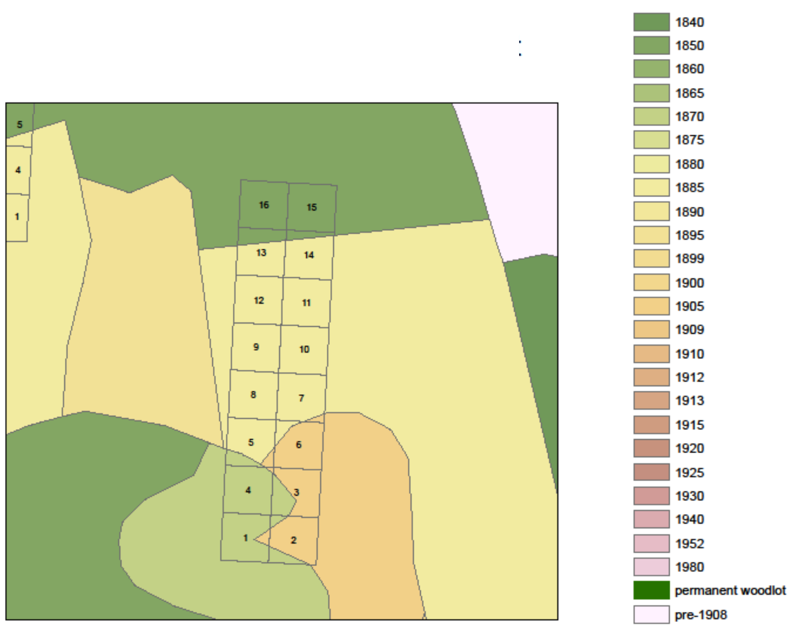
Appendix Figure 11. Prospect Hill Plot 7 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



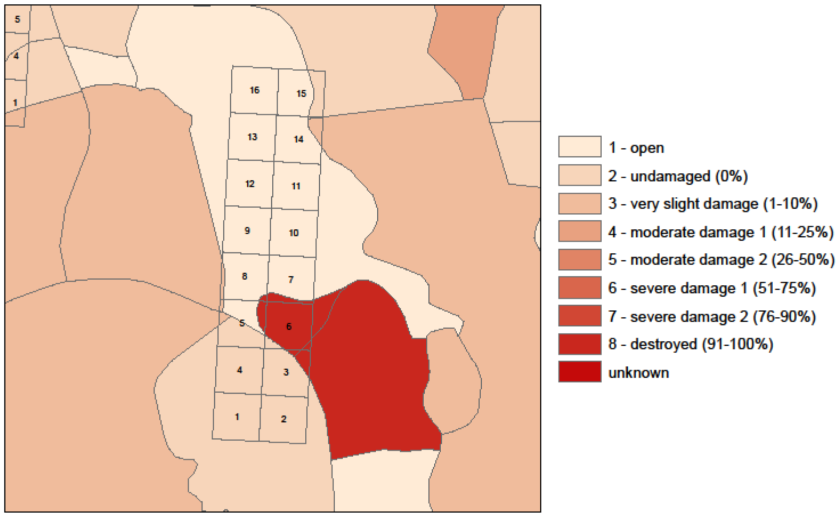
Appendix Figure 12. Prospect Hill Plot 7 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



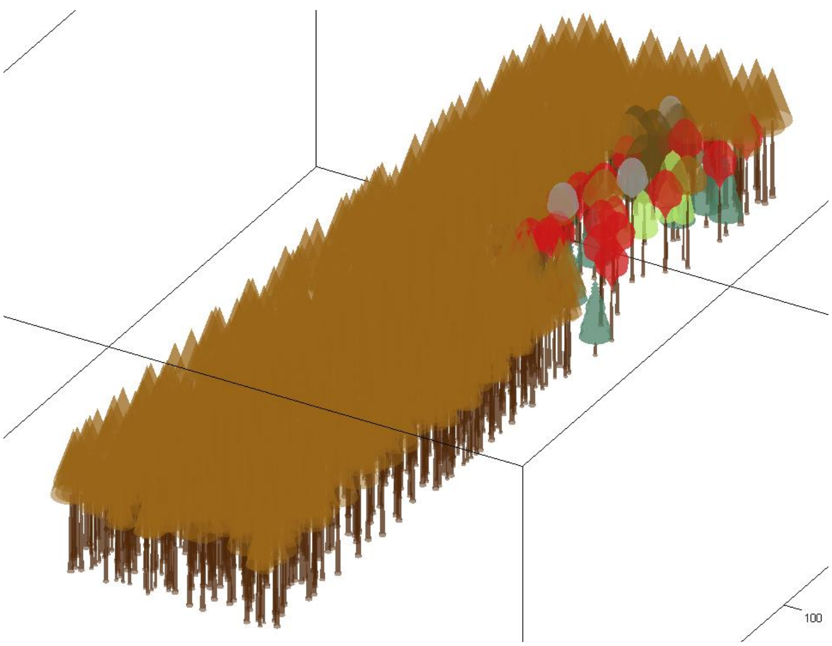
Appendix Figure 20. Prospect Hill Plot 9 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



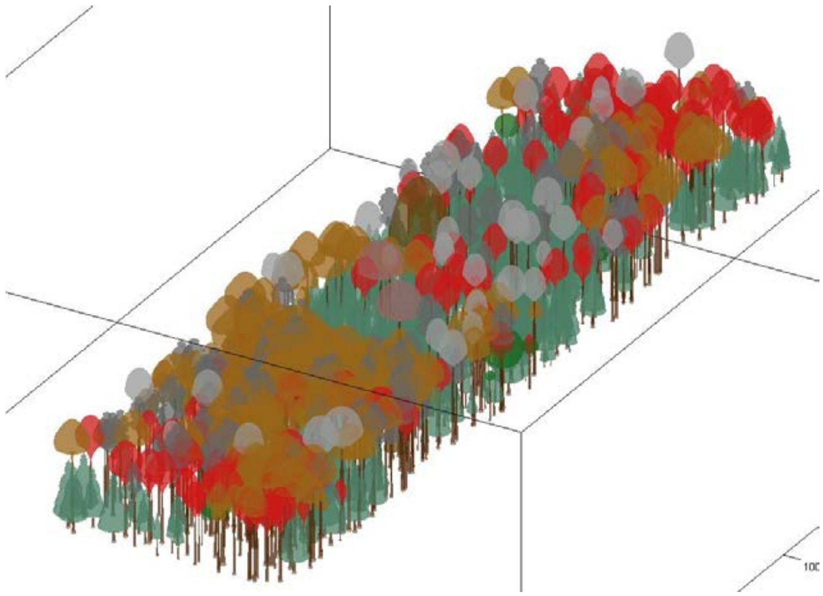
Appendix Figure 21. Prospect Hill Plot 10 Agricultural Abandonment. Older ages of abandonment are shown by darker green with more recent abandonment shown as tan to brown.



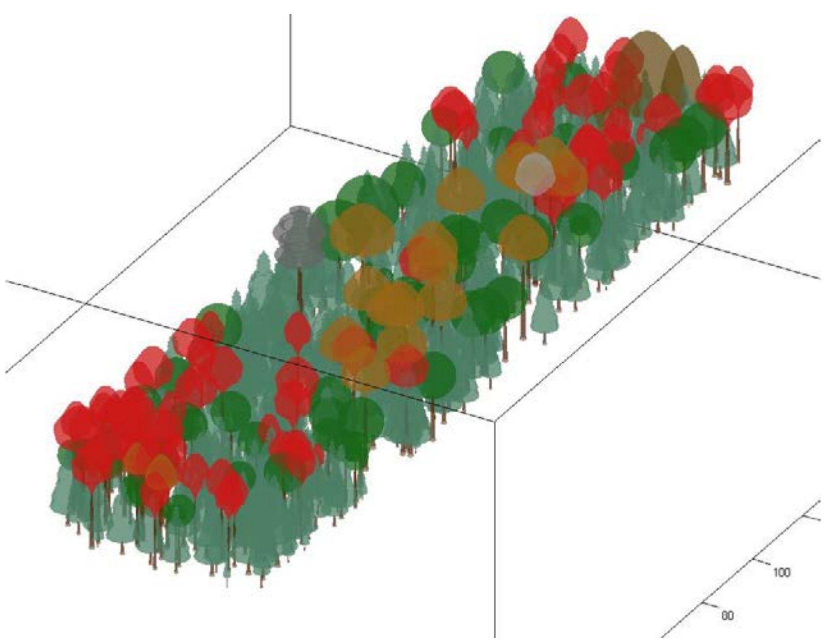
Appendix Figure 22. Prospect Hill Plot 10 1938 Hurricane Damage. Higher levels of damage are indicated by darker red shades.



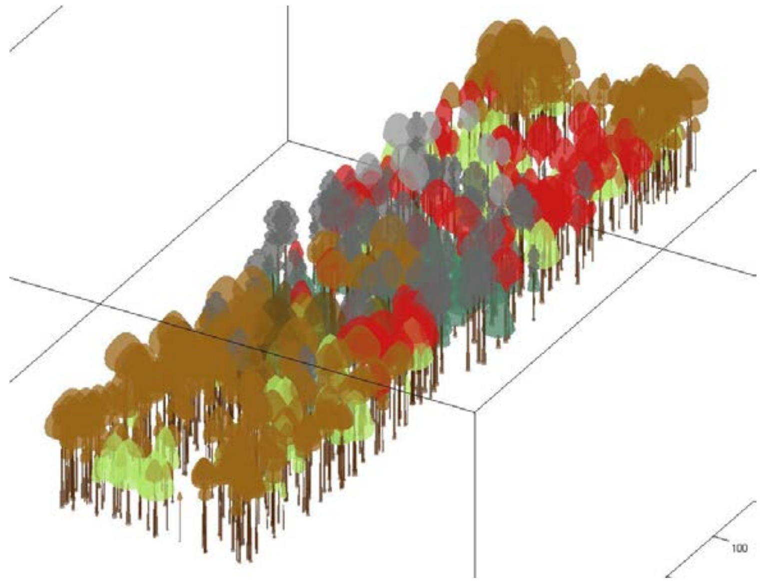
Appendix Figure 23. Visualization of Prospect Hill Plot 1.



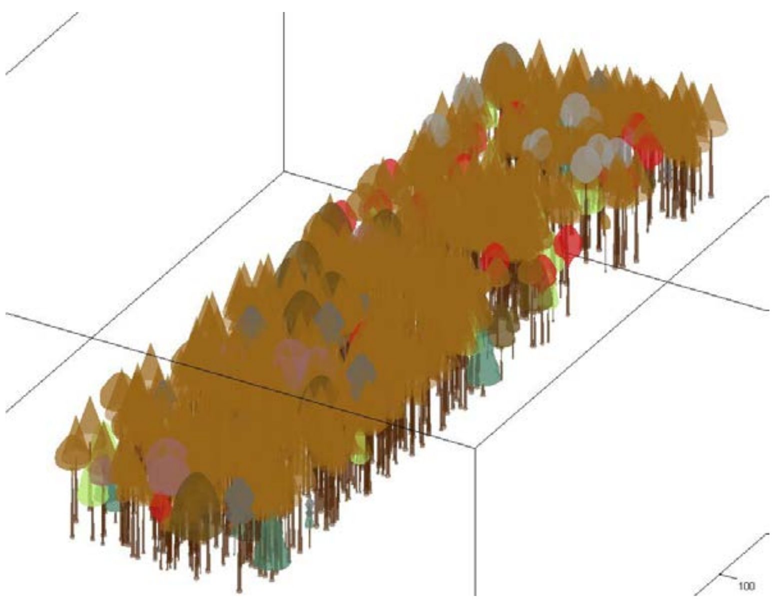
Appendix Figure 24. Visualization of Prospect Hill Plot 2.



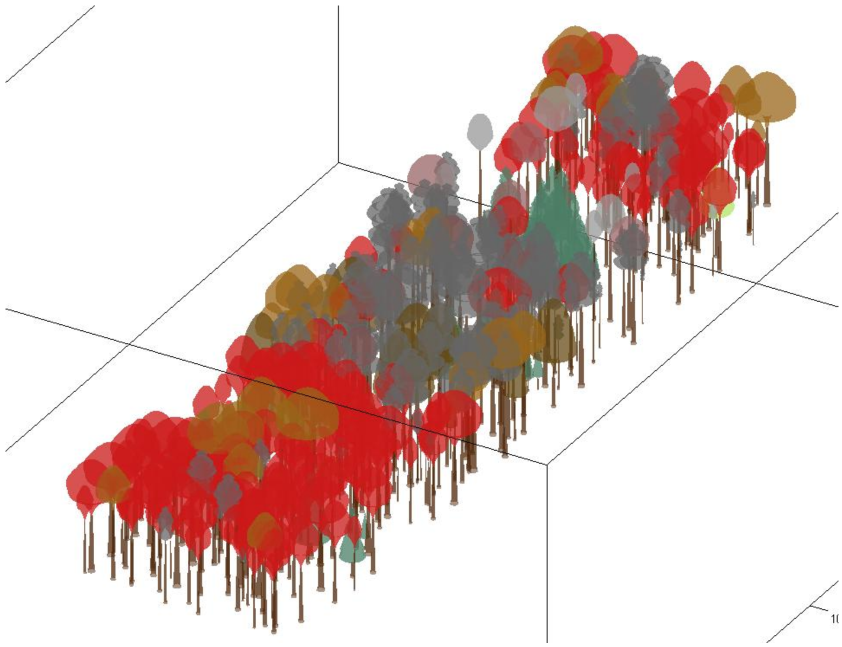
Appendix Figure 25. Visualization of Prospect Hill Plot 4.



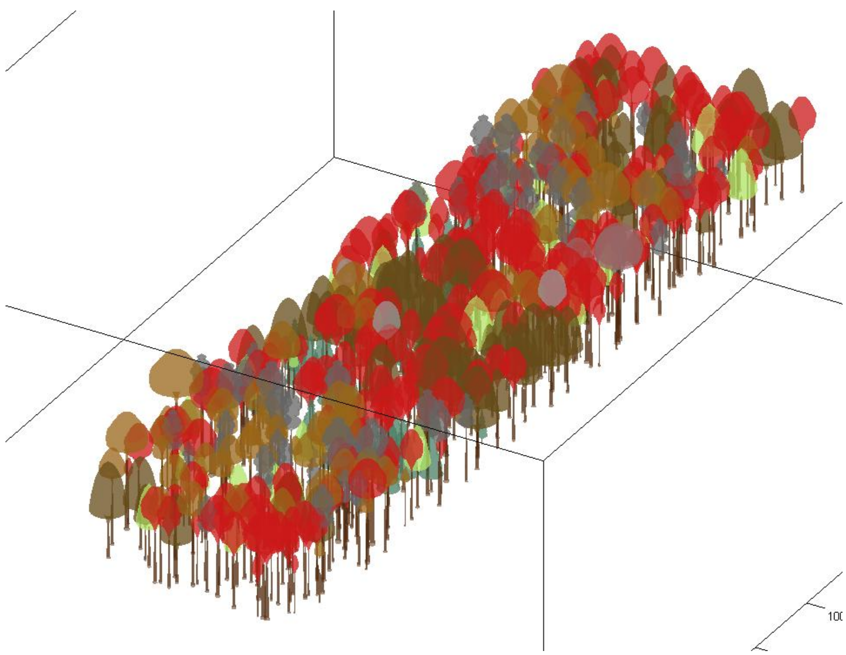
Appendix Figure 26. Visualization of Prospect Hill Plot 5.



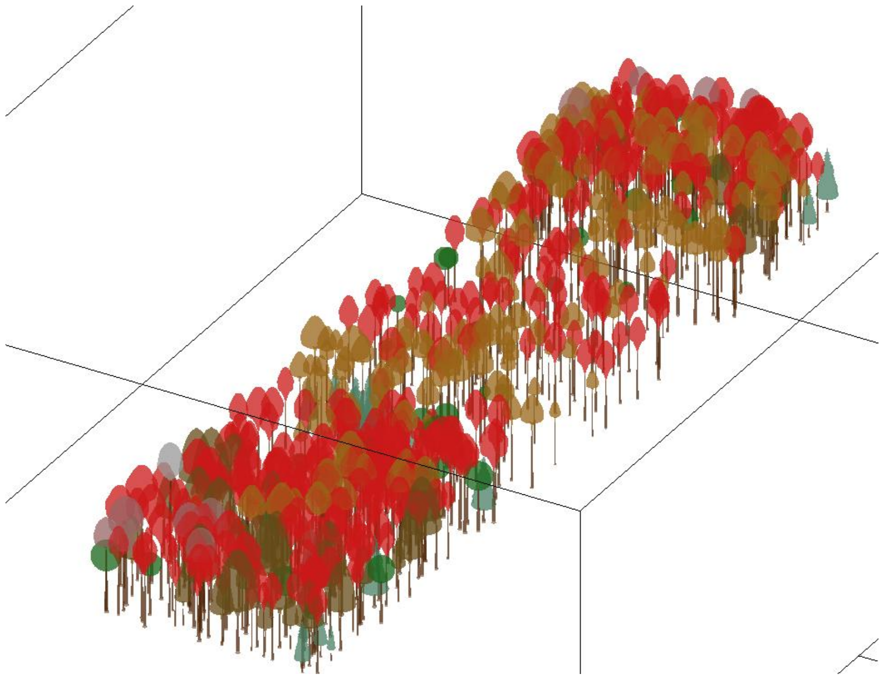
Appendix Figure 27. Visualization of Prospect Hill Plot 7.



Appendix Figure 28. Visualization of Prospect Hill Plot 8.



Appendix Figure 29. Visualization of Prospect Hill Plot 9.



Appendix Figure 30. Visualization of Prospect Hill Plot 10.

10. APPENDIX 2: SORTIE PROCESSING SUITE CODE

```
function[gz] = batch_gunzip(directory,outputdir)
% function [gz] = batch_gunzip(directory,outputdir)
% gunzip(files) uncompresses all GNU zip files from the specified
% directory. Output files have the same name, exluding the .gz and are
% stored in the specified output directory.



if(~strcmp(directory(end),filesep))
filesearch = [directory filesep'*28.xml.gz'];
% '*xx.xml.gz' specifies the timestep you wish to retrieve
else
filesearch = [directory '*28.xml.gz'];
directory = directory(1:(end-1));
end

D = dir(filesearch);
nfiles = length(D);


if(nfiles == 0)
fprintf('No gz files found in the directory: %s\n',directory);
return
end

gunzip([directory filesep D(1).name],outputdir);
forifile = 2:nfiles,
gunzip([directory filesep D(ifile).name],outputdir)
end
end
```

```
function[tar] = batch_untar(directory,outputdir)
%function [tar] = batch_untar(directory,outputdir)
%extracts contents of all tarball archives within a folder into the
specified outputdir
%directory.
if(~strcmp(directory(end),filesep))
filesearch = [directory filesep'*.tar'];
else
filesearch = [directory '*.tar'];
```

```
directory = directory(1:(end-1));
end

D = dir(filesearch);
nfiles = length(D);


if(nfiles == 0)
fprintf('No tar files found in the directory: %s\n',directory);
return
end

untar([directory filesep D(1).name],outputdir);
forifile = 2:nfiles,
untar([directory filesep D(ifile).name],outputdir)
end
end




function tree = sortie_read(file_name)
% function tree = sortie_read(file_name)
%
% read an xml file that has been generated by the SORTIE program.  The
% output is a matlab structure which contains information about
individual
% trees in the simulation.example filename:  'PH3large.xml'

% xdoc = xmlread(which(file_name))
xdoc = xmlread(file_name);

% first, get all of the species names
tm_species = xdoc.getElementsByTagName('tm_species');
Nspecies = tm_species.getLength;
forispecies = 1:Nspecies
species_name(ispecies) = {char(tm_species.item(ispecies-
1).getAttribute('speciesName'))};
end

% this next set of routines is not necessary
if(false)
% then, find the "tp" code associated with the species names
tm_tree_settings = xdoc.getElementsByTagName('tm_treeSettings');
Nspecies = tm_tree_settings.getLength;
forispecies = 1:Nspecies
species_name(ispecies) = {char(tm_tree_settings.item(ispecies-
1).getAttribute('sp'))};
tp_code(ispecies) = str2num(char(tm_tree_settings.item(ispecies-
1).getAttribute('tp')));
end
end

% relate species names to the common names of trees
```

```
forispecies = 1:Nspecies
switch char(species_name(ispecies))
case'ACRU'
common_name(ispecies) = {'Red Maple'};
case'ACSA'
common_name(ispecies) = {'Sugar Maple'};
case'BEAL'
common_name(ispecies) = {'Yellow Birch'};
case'FAGR'
common_name(ispecies) = {'American Beech'};
case'TSCA'
common_name(ispecies) = {'Eastern Hemlock'};
case'FRAM'
common_name(ispecies) = {'White Ash'};
case'PIST'
common_name(ispecies) = {'White Pine'};
case'PRSE'
common_name(ispecies) = {'Black Cherry'};
case'QURU'
common_name(ispecies) = {'Red Oak'};
case'PIRE'
common_name(ispecies) = {'Red Pine'};
otherwise
common_name(ispecies) = {'unknown'};

end
end

trees = xdoc.getElementsByTagName('tree');
Ntrees = trees.getLength;
foritree = 1:Ntrees
tree_characteristics = trees.item(itree-1);
% the values of 'x' and 'y' are swapped below from the definition in
the xml file
tree(itree).y = str2num(tree_characteristics.item(0).getTextContent);
tree(itree).x = str2num(tree_characteristics.item(1).getTextContent);
tree(itree).dbh = str2num(tree_characteristics.item(2).getTextContent);
tree(itree).hv = str2num(tree_characteristics.item(3).getTextContent);
tree(itree).crown_radius =
str2num(tree_characteristics.item(4).getTextContent);
tree(itree).crown_depth =
str2num(tree_characteristics.item(5).getTextContent);
tree(itree).biomass =
str2num(tree_characteristics.item(6).getTextContent);
tree(itree).species = char(species_name(1 + str2num(trees.item(itree-
1).getAttribute('sp'))));
tree(itree).common = char(common_name(1 + str2num(trees.item(itree-
1).getAttribute('sp'))));
tree(itree).sp_code = str2num(trees.item(itree-1).getAttribute('sp'));
end

return
```

```
function stats = sortie_subplot_stats(tree,Dx,Dy,overlap)
% function stats = sortie_subplot_stats(tree,Dx,Dy,overlap)
%
% compute statistics of the input file on a basis of Dx x Dy m
subplots.
% If Dx&Dy are not specified, they are assigned to be 25m each.
Statistics
% are computed with no overlap in either direction; this can be changed
by
% specifying the overlap in the function call to be something else
(e.g. 12.5)

x = flatten(tree.x);   % the x & y coordinates
y = flatten(tree.y);

min_delta = 5;   % minimum increment in spatial dimension (used for
truncation)

xmax = ceil(max(x)/min_delta)*min_delta;
ymax = ceil(max(y)/min_delta)*min_delta;
xmin = floor(min(x)/min_delta)*min_delta;
ymin = floor(min(y)/min_delta)*min_delta;

if(~exist('Dx','var'))
Dx = 25;
end

if(~exist('Dy','var'))
Dy = 25;
end

if(~exist('overlap','var'))
overlap = 0;
end

% overlap = 12.5;
% overlap = 0;
% delta_side = 25;   % the size of subplot
step_size_x = Dx - overlap;
step_size_y = Dy - overlap;
area = Dx*Dy / (100*100);   % portion of one hectare of the subplot

ii = 0;
jj = 0;
for xx = xmin:step_size_x:(xmax-Dx);
foryy = ymin:step_size_y:(ymax-Dy);
idex = find( (x >= xx) & (x <xx+Dx) & (y >= yy) & (y <yy+Dy) );
% fprintf('sortie_subplot_stats:  ii=%d
length(idex)=%d\n',ii,length(idex));
if(length(idex)~=0)
ii = ii + 1;
```

```
stats = sortie_stats(tree(idex),area);
N(ii) = stats.N;
biomass(ii) = stats.biomass;
rh100(ii) = stats.rh100;
density(ii) = stats.density;
else
jj = jj + 1;
end
end
end

if(jj~=0)
fprintf('  Note:  there were %d subplots with zero trees inside.  This
effect is not \n   included in the statistics\n',jj);
end

clearstats
stats.biomass = biomass;
stats.density = density;
stats.rh100 = rh100;

return


function h = make_tree(tree)
% function h = make_tree(tree)
%
%

trunk_height = tree.hv - tree.crown_depth;
x = tree.x;
y = tree.y;
trunk_profile =
[1.6,1.2,1.15,1.11,1.1,1.05,1,0.99,0.95,0.92,0.91,0.85,0.8,0.75,0.7,0.6
5,0.6,0.55,0.5];

[xt,yt,zt] = cylinder((trunk_profile*tree.dbh)/100,24);   % dbh is in
units of cm, 24 points in circumference
zt = zt * trunk_height;

if(top_level)
clf;
holdon;
end

trunk_standard = [80 40 10]/255;
alpha_standard = 0.5;
switchtree.species
case'ACRU'% red maple
crown_shape = 'profile';
crown_color = [.8 .1 .1];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
```

```
crown_profile =
[0.4,0.8,1.4,2.0,2.9,3.8,4.8,6.0,7.0,8.6,12.4,13.0,13.1,13.3,13.8,14.0,
13.8,13.0,12.0,11.0,9.8,8.8,8.0,6.5,5.0,0];
crown_profile = crown_profile/max(crown_profile);
case'ACSA'% sugar maple
crown_shape = 'ellipsoid';
crown_color = [.4 .7 .1];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
case'BEAL'% yellow birch
crown_shape = 'profile';
crown_color = [.7 .9 .4];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[1.0,10.0,11.3,11.2,11.0,10.5,10.0,9.6,9.1,8.5,8.1,7.8,7.2,6.6,5.0,4.8,
3.2,1.3,0];
crown_profile = crown_profile/max(crown_profile);
case'FAGR'% american beech
crown_shape = 'profile';
crown_color = [.4 .3 .1];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[1.0,1.8,3.0,13.0,12.8,12.2,11.8,11.2,10.5,9.8,8.9,7.5,6.0,4.0,0];
crown_profile = crown_profile/max(crown_profile);
case'TSCA'% hemlock
crown_shape = 'profile';
crown_color = [.3 .5 .4];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[1.0,4.0,8.0,11.0,9.0,10.1,8.8,9.8,10.0,8.1,8.9,8.3,7.9,7.0,7.3,5.2,6.4
,5.0,6.6,6.4,6.2,4.0,5.5,3.2,4.2,3.0,1.9,3.0,2.0,2.8,0.9,2.1,0.9,0.3,0]
;
crown_profile = crown_profile/max(crown_profile);
case'FRAM'% white ash
crown_shape = 'profile';
crown_color = [.6 .6 .6];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[1.0,2.3,4.0,5.7,7.0,8.1,9.2,10.3,11.1,11.0,10.8,10.4,10.2,9.8,9.5,9.0,
8.2,7.2,6.2,4.3,0];
crown_profile = crown_profile/max(crown_profile);
case'PIST'% white pine
crown_shape = 'profile';
crown_color = [.4 .4 .4];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
```

```
trunk_alpha = alpha_standard;
crown_profile =
[1.0,2.0,3.2,4.2,4.9,5.2,4.3,3.0,1.8,3.0,4.3,4.3,3.3,4.0,5.5,5.2,4.2,3.
2,4.3,3.8,2.9,1.2,1.8,2.2,1.0,0];
crown_profile = crown_profile/max(crown_profile);
case'PRSE'% black cherry
crown_shape = 'profile';
crown_color = [.6 .4 .4];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[1.0,1.3,2.0,2.6,3.6,12.4,12.6,12.9,13.0,13.1,13.0,12.5,12.0,11.6,11.0,
10.0,9.0,7.6,5.8,3.5,0];
crown_profile = crown_profile/max(crown_profile);
case'QURU'% red oak
crown_shape = 'profile';
crown_color = [.6 .4 .1];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
crown_profile =
[0.8,0.9,1.0,1.3,2.0,3.0,5.2,14.0,14.2,14.3,14.2,14.1,13.5,12.2,11.0,10
.0,8.7,7.4,6,4,0];
crown_profile = crown_profile/max(crown_profile);
otherwise% unknown
crown_shape = 'sphere';
crown_color = [.1 .4 .1];
crown_alpha = alpha_standard;
trunk_color = trunk_standard;
trunk_alpha = alpha_standard;
end

ht = surf(xt+x,yt+y,zt);
set(ht,'edgecolor','none');
set(ht,'facecolor',trunk_color);
alpha(ht,trunk_alpha);

switchcrown_shape
case'cone'% a cone shape
    [xc,yc,zc] = cylinder(tree.crown_radius,24);
xc(2,:) = 0;
yc(2,:) = 0;
zc(2,:) = tree.crown_depth;
case'ellipsoid'
    [xc,yc,zc] =
ellipsoid(0,0,0,tree.crown_radius,tree.crown_radius,tree.crown_depth/2)
;
zc = zc + tree.crown_depth/2;
case'cylinder'
    [xc,yc,zc] = cylinder(tree.crown_radius,24);
case'profile'
    [xc,yc,zc] = cylinder(crown_profile*tree.crown_radius,24);
zc = zc*tree.crown_depth;
```

```
otherwise% a sphere;
    [xc,yc,zc] = sphere;
xc = xc*tree.crown_radius;
yc = yc*tree.crown_radius;
zc = zc*tree.crown_radius + tree.crown_radius;
end
hc = surf(xc+x,yc+y,zc+trunk_height);
set(hc,'edgecolor','none');
set(hc,'facecolor',crown_color);
alpha(hc,crown_alpha);

if(top_level)
holdoff;
end

h = [hthc];

return
```