SPATIAL ANALYSIS OF OAK WILT SPREAD IN WESTERN NORTH CAROLINA

by

Sharon Yeh

Date:_____

Approved:

_____

Dr. Patrick Halpin, Advisor

Masters project submitted in partial fulfillment of the requirements for the joint Master of Forestry and Master of Environmental Management degree in the Nicholas School of the Environment and Earth Sciences of Duke University

2007

TABLE OF CONTENTS

ABSTRACT

Oak wilt, *Ceratocystis fagacearum*, has been documented in the Southern Appalachian region of North Carolina since the early 1950s. Due to its rapid spread and rate of damage, the North Carolina Forest Resources Division monitored oak wilt closely by performing annual surveys to both control for and monitor the disease. This project uses the monitoring data to investigate where and how oak wilt could spread in the near future. Understanding the spatial nature of oak wilt can help managers target future monitoring and prevention efforts for this particular region. Both spatial and statistical analyses were used, including the Chi-Squared test, Classification and Regression Tree (CART) models, and the Mantel test. Results from several tests indicate that oak wilt prefers specific oak species within the red oak family, has the potential to spread in Western North Carolina, and spread of oak wilt by long-range pathogen mechanisms have a higher impact on the transmission of oak wilt than short-range mechanisms.

INTRODUCTION

Valued for timber and food production, genus *Quercus* account for approximately 35% of all hardwood volume in the country. Oak wilt, in conjunction with other diseases such as sudden oak death, poses a serious threat to our current forest composition. In fact, oak wilt has been recognized as one of the most destructive diseases to affect the oak species in the country (Wilson, 2001). First officially identified in the 1940s in Wisconsin, oak wilt was likely observed earlier by loggers noting dying oaks in the upper Mississippi River Valley as early as the late 1800s to the early 1900s (Gibbs and French, 1980). Originating in eastern Russia, this disease was probably introduced to North America by trade. Oak wilt has since dispersed to its current 22 state range, affecting much of the central and eastern U.S., and reaches as far as Texas.

Majority of oak species are categorized within two subgenera; red or black oaks (*Quercus* subgenera *Erythrobalanus*) or white oaks (*Quercus* subgenera *Quercus).* Oak wilt is a disease that causes damage and loss primarily across the red oak family group. It is caused by a fungus, *Ceratocystis fagacearum,* and spreads through root grafts, via sap-feeding insects, and via tree-wounding insects such as bark beetles. It was feared that oak wilt can be devastating to forests because it can infect and cause oaks within the red oak family to die within a year of infection, and host populations are enormous within the known range of the disease (Hepting, 1955).

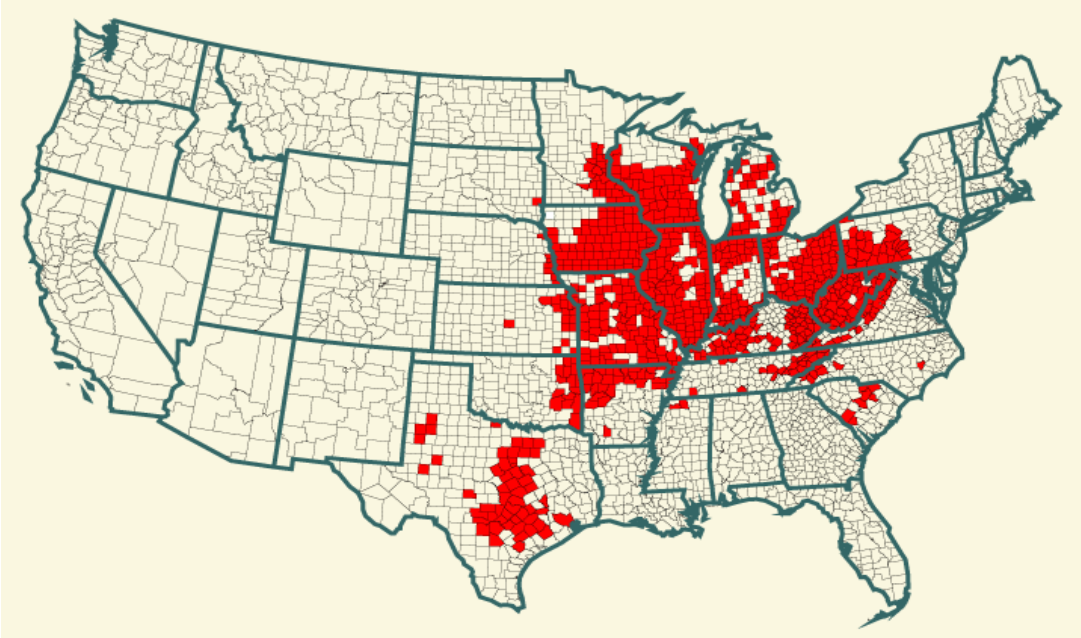


Figure 1 – Oak wilt distribution in the United States, 2005, USDA Forest Service.

This disease was first discovered in 1951, in the mountains of Haywood and Buncombe County. The North Carolina Forest Resources Division responded by initiating an annual oak wilt survey within the same year, as potential threat and severity of the disease was unknown. Aerial surveys were followed with fieldwork on the ground so that field crews could 1) take a culture of the specimen, 2) record information on the diseased oak species, 3) and treat the affected trees. Through the annual surveys, NC Forest Service has been able to contain most of the diseased trees and prevent massive outbreaks. Since the focus of the surveys has been towards containment, much of the collected data has not been analyzed regarding predictive aspects of oak wilt. Questions regarding species preference, ecological indicators, and spatial distribution still remain unknown.

Although literature on oak wilt in this region was abundant when it first appeared, the number of published research declined dramatically by the 1960s. There may be several reasons for this. First, oak wilt spread seemed to be slower than anticipated. An early study of oak wilt in the Southern Appalachian region conducted by Boyce in 1957 showed that activity of wilt centers would naturally cease after an average of three years, if treated. Boyce also found that of 163 infected centers in North Carolina, Kentucky, and Tennessee, ninety-six percent of wilting oaks were found within 50 feet of trees that died of wilt. Of this ninety-six percent, seventy-eight percent were found within 30 feet of trees that died of wilt. The study concluded from this data that disease spread was either due to root grafts or to very short, above ground transmission of the fungus. Because spread seemed to be conducted mostly through root grafting and only some overland flow, wilt was also found mostly in clusters, rather than a dramatic spread of wilt across an area. Thus, with aggressive control and monitoring, states were able to control the spread of wilt quite effectively in this region.

Another study (Liebhold et al., 1995) posed two additional explanations for the slow spread of oak wilt in the Southeast. The rich species diversity in the Southern Appalachian Mountains may cause a greater spatial distances, or separation between oak species. Secondly, this study also suggested that soils in this area might be less suitable for root grafting. On the other hand, Texas is currently experiencing a rapid rate of oak wilt spread predominantly in live oaks, which have extensive root systems that readily spread and graft in sandy soils (Appel, 1995).

Dekker (1978) used the same dataset from NC Division of Forest Resources to analyze the spread patterns and distribution of oak wilt fungus in Buncombe and Haywood County. Completed in 1978, he found that the spread of oak wilt originated from Tennessee, and because the physical barrier of the large Appalachian Mountain range exists between Tennessee and North Carolina, human activity was suspected as the main culprit in transmitting the disease to North Carolina. He also found that oak wilt affected trees were found mostly between 2,000 and 3,000 feet, and that the direction of oak wilt spread was not uniform; it spread eastward in Haywood county, but spread westward in Buncombe county. However, the paper still called for further analysis to determine if oak wilt preference towards the 2,000 to 3,000 ft elevation was caused by a species preference of the red oak group in that elevation gradient, or because the fungus prefers that elevation regardless of species composition.

OBJECTIVES

Because the available literature was based mostly on data collected within the first decade of oak wilt spread, and not many studies were completed regarding wilt spread specifically for the Southern Appalachians, I was interested in analyzing predictive qualities of oak wilt within western North Carolina region using data collected over a longer period of time. In this study, I will analyze the same NC Forest Resources dataset of oak wilt surveys from 1951 to 1997 to determine 1) if oak wilt has a particular species preference among the affected oaks, 2) what the predicted habitat of where oak wilt can occur within the currently affected area, and 3) what factors have the most impact on determining whether oak wilt will occur?

STUDY AREA

These three questions were then applied to a five-county study area in the mountains of western North Carolina. The five counties are Jackson, Swain, Haywood, Buncombe, and Madison, and were chosen because the majority of affected trees were found within their boundaries. These counties are located mostly within the Southern Appalachian Mountain range, known for high biodiversity with patches of old growth forest. The forest composition in this area ranges from mixed hardwoods in the lower elevations to mixed coniferous forest stands in the higher elevations. Metropolitan areas are found mostly in the valley areas, with Asheville being the largest town located within Buncombe County. Some cultivated areas exist, but are not predominant in this mountainous terrain.

MAP OF STUDY AREA

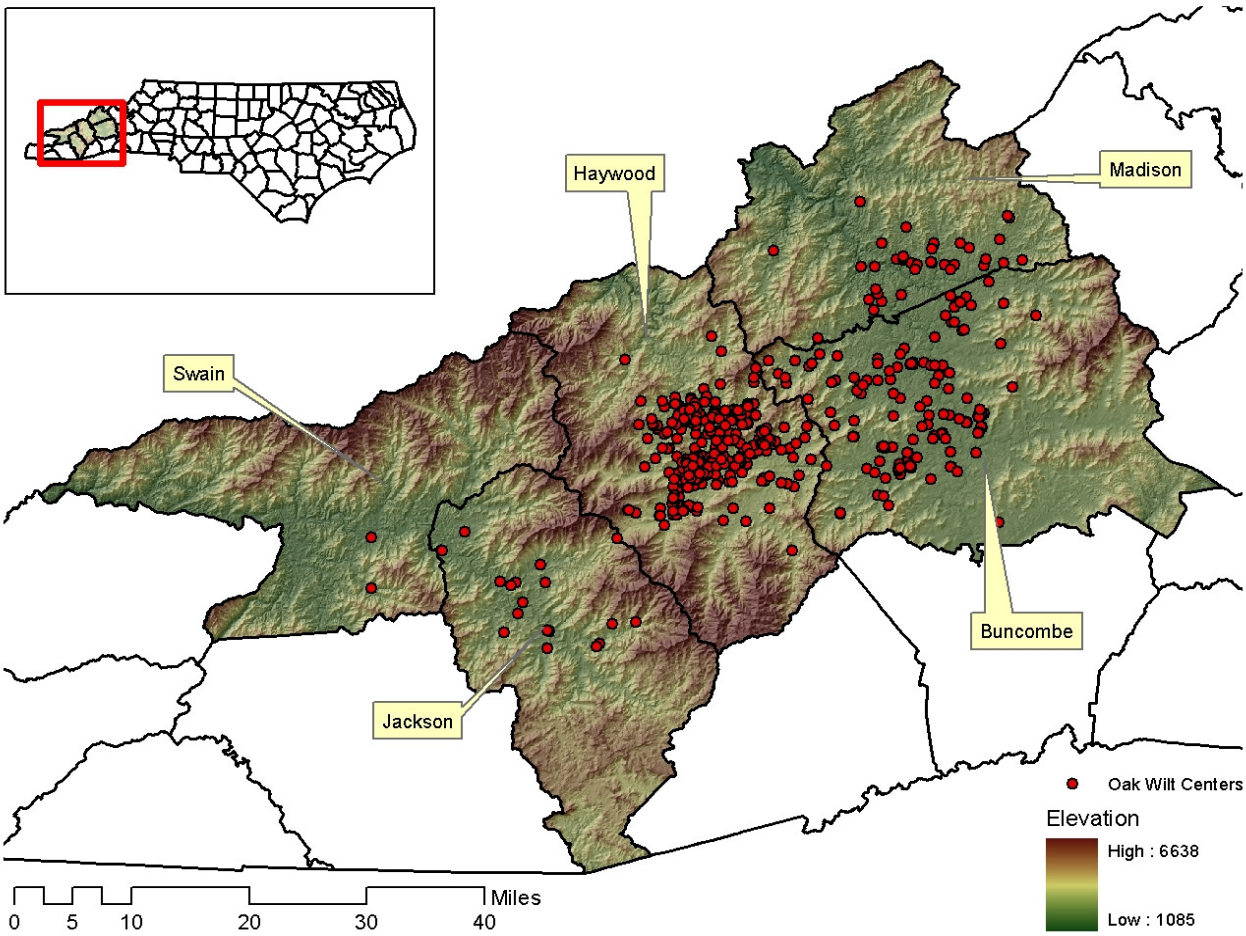


Figure 2 – Study area include the North Carolina counties of Swain, Jackson, Haywood, Madison, and Buncombe. Red dots represent presence of oak wilt centers from 1951-1997.

As Figure 2 shows, many of the wilt centers were concentrated in the valleys of Haywood County, with some clusters in Buncombe and Madison County. Spread to Jackson and Swain County has been limited.

DATA COLLECTION, METHODS, RESULTS, AND DISCUSSION PER OBJECTIVE

Dataset Collection Method

The Pest Management division of the NC Division of Forest Resources supplied the dataset of suspected oak wilt centers, which contained location, tree species, dbh, age of infection, and average size of infection. Data was collected through annual surveys conducted through flight lines along contours, due to the mountainous terrain of the area. Performed at an elevation between 500 and 1000 ft, aerial surveys were conducted annually from June through August, to avoid leaf-fall season. The Pest Control Division was able to achieve full aerial coverage of the five counties during their study. Oaks with discolored foliage were easily spotted from the air and locations were sketched for subsequent ground-truthing and treatment. Prior to 1970, branch samples were taken in the field to create a culture sample following the aerial survey. After 1970, however, an increment core was taken to produce a culture. Before 1970, if a tree was positively identified with oak wilt, it would have been felled and treated with fungicide-insecticide, and stumps would be treated to prevent disease spread through root grafting. After 1970, trees were treated with herbicide only while standing. The theory was to kill the root systems to reduce spread through root grafting, and to dry the tree bole to reduce fungal mat formation.

Additional data collection methods will be described they pertain to each separate objective and method.

Descriptive Statistics

Figure 2 summarizes the average values of several descriptive characteristics regarding oak wilt centers by county. We can see that Haywood County exhibited the highest number of wilt centers, while only two centers were reported in Swain County. Average values across counties did not vary greatly, although the diameter at breast height (dbh) is higher in Swain County with a value of 22.38 inches as compared to the average of around 16 inches for the remaining counties. This is most likely due to the small *n* available for statistical calculations.

County	n	DBH (in)	Elevation (ft)	Distance to Stream (ft)	Distance to Development (ft)	Distance to Cultivated (ft)
Buncombe	157	16.33	2644.63	1748.87	1181.68	1051.99
Haywood	435	16.30	2646.76	1747.15	1185.18	1050.94
Jackson	21	16.51	2655.81	1774.49	1220.31	1070.40
Madison	42	16.21	2646.57	1746.10	1181.42	1048.78
Swain	2	22.38	2651.61	1477.54	1094.79	1039.82
Total Averages	657	16.31	2645.17	1745.58	1182.72	1049.63

Figure 3 – Average values of descriptive characteristics concerning wilt affected trees.

Species	Quercus alba	Quercus coccinea	Quercus falcata	Quercus rubra	Quercus velutina
Count	20	271	32	265	103
Total Number Trees	692				
Percent of total	**3%**	**39%**	**5%**	**38%**	**15%**

Figure 4 – Species distribution wilt affected trees within the North Carolina counties of Haywood, Jackson, Swain, Buncombe, and Madison.

Only 20 infected centers featuring *Quercus alba* were recorded, as compared to 265 counts for *Quercus coccinea*.(Fig. 4).

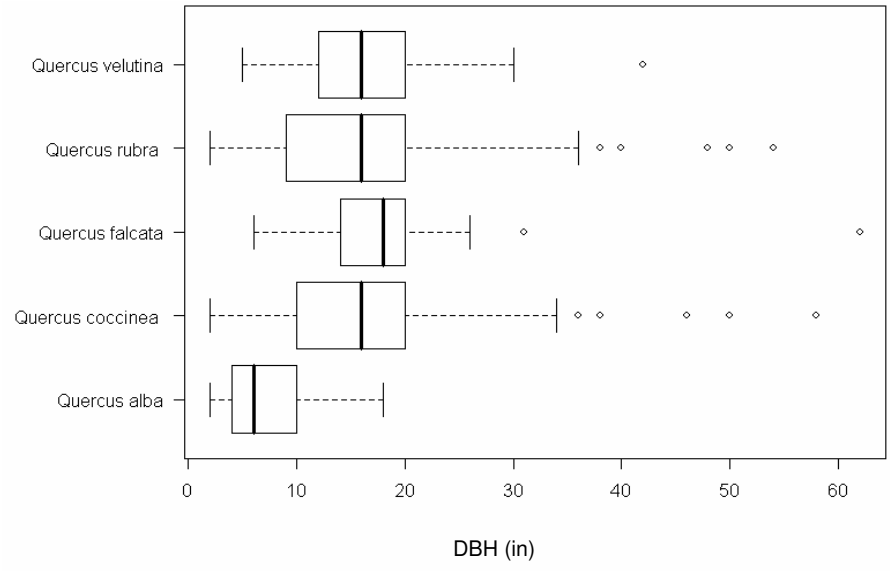


Figure 5 – Boxplot representing the distribution of DBH(in) values, categorized by species.

While the average diameter at breast height (dbh) for all wilt affected trees was 16.31 inches (Fig. 3), the range of values varied greatly by species (Fig. 5). The upper quartile was around 20 inches for all species except *Quercus alba*, while the lower quartile varied by species. Outliers

for *Quercus rubra* and *Quercus coccinea* show the presence wilt in mature trees, while none are reported for *Quercus alba*. This figure also shows that the median size of the tree is much smaller for *Quercus alba* than other species within the red oak family.

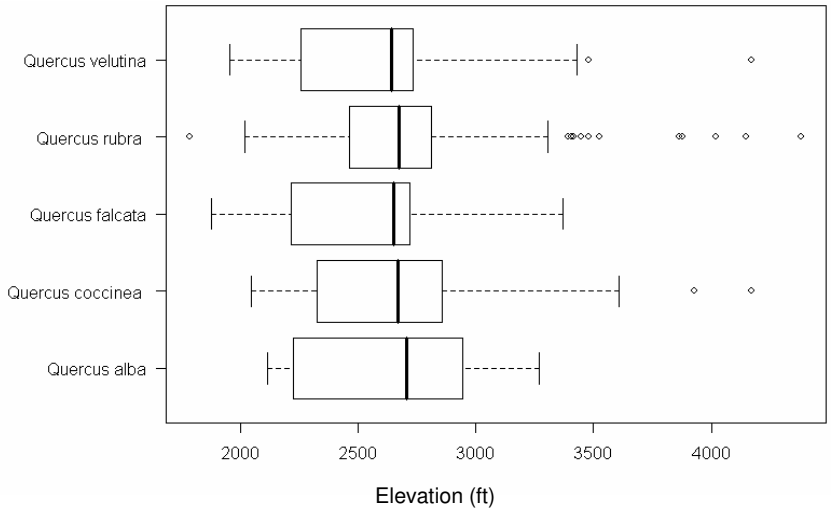


Figure 6 – Boxplot representing the distribution on elevation values for wilt affected trees, categorized by species.

The average elevation of all wilt affected trees is 2,645 ft (Fig. 3). This boxplot shows that the interquartile elevation values across all species lie between 2,000 ft to 3,000 ft, indicating that we can find most of the wilt affected trees within this range (Fig. 6). However, the outliers are numerous in *Quercus rubra*, perhaps indicating that the species distribution spans a larger elevation gradient.

Objective 1 – Chi-Squared Test of Proportions for Species Preference

The first objective of this study asks if oak wilt affects certain oak species more often than others. Determining this could allow forest managers to target specific species in monitoring or treatment of hardwood stands. It can also advise landowners on which oak species to manage for in order to avoid wilt.

Data Collection and Methods: Chi-Squared Test

Although it has been documented that oak wilt prefers the red oak group species in the study area (Rexrode and Brown, 1983), information remains unknown regarding which specific oak species are preferred within this group. To determine this, I performed a Chi-squared test to compare proportions of affected oak species versus the proportions of oak species found in the five county area. This population data was provided by the USFS Forest Inventory Analysis (FIA) group, in the form of total number and volume of live trees within the oak species located in the study area. FIA population data was available for 1984, 1990, and 2002. While this inventory does not include the number of trees found within the National Park Service (NPS) boundaries, NC Division of Forest Resources also excluded NPS lands from the surveys.

Because the dataset provided by NC Division of Forest Resources spans almost fifty years, I was also interested to see if species preference changed over time. I subsequently separated the dataset into two blocks, from 1951-1983 to compare with the 1984 FIA data, and from 1984-1997 to compare with 2002 FIA data.

Contingency Tables

A table that classifies observations in two or more ways is called a contingency table, and reveals how variables are related to one another. For this project, it shows the relationship between oak species and presence of oak wilt as compared to the number of oak species in the population. Contingency tables allow us to test whether these classifications are independent with the following hypothesis:

H_O = *Proportions of affected oak species will reflect the same proportions found in the general population of oak species.*

H_a = H_O *is false*

Chi-Squared Test

The Chi-Squared test of independence can be applied to the contingency table to determine statistical significance in this analysis. These mathematical formulas compare the observed frequencies of a phenomenon with the frequencies we would expect if there were no relationship between two variables in a larger sampled population. For this project, the Chi-Squared tests compare the proportions of oak wilt by species vs. proportions of oak species found in the population. Essentially, it tests actual results against the null hypothesis to assess whether the results are different enough to overcome random sampling (DeGroot, 1986).

Chi-Squared test statistic can be computed as follows:

$$X^2 Obs = \sum_{cells} \frac{(O-E)^2}{E}$$

where O represents observed values, and E represents expected values.

For this project, I used R (R Development Core Team, 2005), a freeware statistical package. The "prop.test" command allows users to test the null hypothesis that proportions of wilt in several groups are the same, and produces a Chi-Squared test statistic to measure the significance of the proportions. It assumes a two-sided distribution with a 95% confidence level.

If the proportions of affected oak species are significantly different from the proportions of oak species found in the five county area, and if they are different enough to overcome random sampling, then some inferences can be made on whether a certain species is preferred or not.

Results and Discussion: Chi-Squared Test

The results of the Chi-Squared tests show that we can reject the null hypothesis and accept the alternative hypothesis for both time periods. In the first time period, from 1951-1983, the difference between the proportions of wilt affected trees by species vs. the proportions of species found within the population are statistically significant, with a Chi-Squared test statistic (X^2) of 585.45 and p-value of 2.20E-16 (Fig. 7).

1951-1983 vs. 1984 FIA data Chi-Squared Test: Proportions of species within study area						
Species	**Wilt affected**	**n**	**% found in population (FIA)**	**n**	**p-value**	**Chi-Squared test statistic (X^2)**
Quercus alba	1.16%	6	33.76%	26,114,001	2.20E-16	585.45
Quercus coccinea	45.16%	233	22.64%	17,512,994		
Quercus falcata	4.46%	23	0.48%	368,700		
Quercus rubra	32.36%	16	24.46%	18,923,996		
Quercus velutina	16.86%	87	18.67%	14,440,001		
Totals		365		77,359,692		

Figure 7 – Summary table of 1951-1983 vs. 1984 FIA Chi-Squared Test

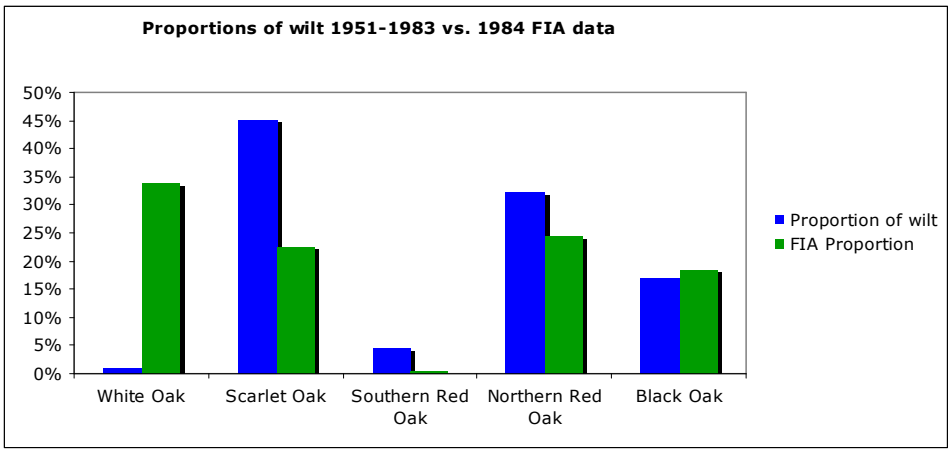


Figure 8 – Graph representing oak wilt's species preference from 1951-1983

The graph above (Fig. 8) reveals that oak wilt did prefer some species over others in the first time period. Wilt strongly prefers *Quercus coccinea* (scarlet oak), representing 45% of the wilt-affected trees, while representing 22% of the total population. There is also a slight preference for *Quercus rubra* (northern red oak), which represented 32% of the affected trees, while representing 24% of the total population. Avoidance of oak wilt by *Quercus alba* (white oak), can also be seen, as the affected proportions are significantly lower (1%) than the proportion of white oaks in the study area (33%) (Fig. 7). While no *Quercus* species is immune to infection by oak wilt, the two subgenera are not equally susceptible to infection (Henry et al, 1947). One study suggests that fungal mats, a reproductive structure of oak wilt and an important source of pathogen transmission by insects, are only produced on the red oak family. Because of this, the mating system between trees may influence whether oak wilt can be developed between two different subgenres of oak wilt (Engelhard, 1956).

1984-1997 vs. 2002 FIA data Chi-Squared Test Proportions within study area						
Species	**Wilt affected**	**n**	**% found in population (FIA)**	**n**	**p-value**	**Chi-Squared test statistic**
Quercus alba	22.82%	8	4.55%	6,633,996	2.90E-14	72.6305
Quercus coccinea	23.80%	44	25.00%	6,917,991		
Quercus falcata	3.55%	9	5.11%	1,032,033		
Quercus rubra	35.33%	99	56.25%	10,268,981		
Quercus velutina	14.49%	16	9.09%	4,212,001		
Totals		176		29,065,002		

Figure 9 - Summary table of 1984-1997 vs. 2002 FIA Chi-Squared Test

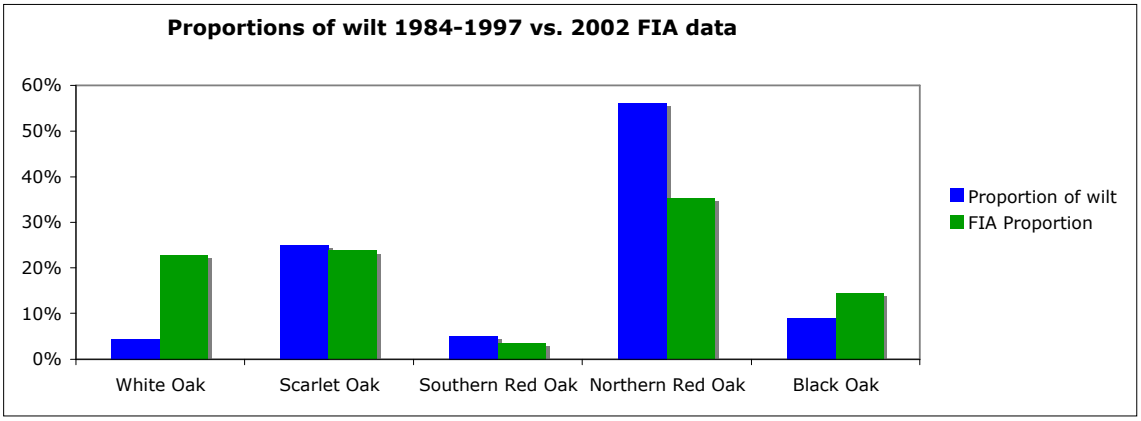


Figure 10 - Graph representing oak wilt's species preference from 1984-1997

In the second time period, we can see the strong species preference moving from scarlet oak to northern red oak. The total number oaks found in the study area decreased dramatically by 48 million trees. The total number of northern red oak also decreased in this time period by approximately 8 million trees, while the proportion of northern red oak among all oak species increased from approximately 24% to 55% (Fig. 9). In addition, the total number of scarlet oak decreased by approximately 10 million trees, while the proportion of scarlet oak found in the environment increased slightly from 22% to 25% of oak species.

With the Chi-Squared test, we can conclude that there was indeed a species preference for oak wilt. Knowing that northern red oak is more susceptible to oak wilt, landowners with hardwood stands may want to manage for other species such *Quercus alba*, or white oak. While these results show a shift in wilt preference for northern red oak from scarlet oak, the reliability of the FIA population data cannot be guaranteed. Some error in species identification could have occurred as these two species are hard to differentiate from each other, and field crews have

limited training before their assignments. While FIA data is the best available data we have regarding the population, it is still only a predictive model of species composition based permanent sample plots. In addition, other factors external factors to forest composition such as development of climate change may have also impacted species representation.

Please refer to Appendix A for the complete Chi-squared results.

Objective 2 – Predict possible oak wilt habitat through CART Models

While determining the species preference of wilt helps us predict where wilt may occur, geographic datasets showing oak distribution by species have yet to be accomplished for the study area. Therefore, the second objective of this project is to create a predictive model, using environmental variables, to see where oak wilt can potentially occur in the study area. This information will helpfully help NC Division of Forest Resources in managing for this disease.

Data Collection and Methods: CART model

Classification and Regression Tree (CART) models were developed by Breiman et al (1984) as a departure from standard general linear regression models to determine the relationship between response and predictor variables. CART models are used to distinguish differences among groups, and for this particular project, the groups represent presence and absence of oak wilt. Because we are interested in categorical responses, we will be using the classification tree for this project.

There are several features that make this technique applicable for ecological data. CART models are nonparametric and therefore do not assume any particular distribution. They allow for both categorical and numeric predictor variables, and provide easy visual interpretation of explanatory variables (De'ath and Fabricius, 2000). CART models act to recursively partition predictive environmental variables. For each split, the maximum deviance is chosen from the response variable, ultimately creating relatively homogenous groups found at the bottom of the tree (Vayssières et al., 2000).

Explanatory variables used for this study included elevation, aspect, slope, topographic relative moisture index (TRMI) (Parker, 1982), land use cover type, as well as distance from streams, developed areas, and cultivated areas. TRMI combines properties of slope and aspect to create a single index that identifies which sites could feature xeric or mesic environments. For details on the TRMI model, please refer to Appendix B for the full Python script and graphic of the model.

Slope and aspect were derived from a digital elevation map (DEM), which was based off the United States Geological Survey (USGS) DEM 90m cells, and are made available from the

North Carolina Department of Transportation. Land use and land cover (LULC) data was provided by USGS, and organized into 1:250,000 quadrangle tiles. Distances associated with individual wilt points were calculated using Euclidean distances. To create slope, aspect, and TRMI, I used spatial analysis tools in ArcGIS (9.1) to build a model.

Although soil could potentially be an influential explanatory variable due to the possibility of wilt spread through root grafting, the data was insufficient to do an in-depth analysis. While State Soil Geographic .STATSGO) Database was available for the state, the more detailed Soil Survey Geographic (SSURGO) Database was only available two counties, and often created from forest cover data. When initial analysis was done using the SSURGO data, it still didn't provide enough detail as to distinguish past the family level of soil types.

Creating the tree

I first created a model using the "tree" function within the "tree" library found in the statistical package, R. After the initial tree was created, I pruned the tree using the "prune.tree" function since some of the data was over fitted. I then cross-validated the data to assess how well the tree was fitted using the "cv.tree" function. Finally, I created a confusion matrix to determine how well the model classified samples using the "predict.tree" function. After one model was created using the "tree" function, I created another CART model using the "rpart" function, which provides some more detail in explaining how trees were fitted. It shows which predictor variables could have been alternate splits in the tree.

Using the model to create a predictive map

After these models were created, I made a predictive map by plotting the model results in ArcGIS (9.1). I wrote a docell script describing the leaves of the CART model to show predictive groups of geospatial variables. Finally, I used the "predict.tree" function in R to create a table of predicted group memberships versus actual membership values. This function creates a confusion matrix, which examines exactly how samples were misclassified. The overall success rate can be calculated by taking the sum of the diagonal and then taking it over the total number of samples, while the misclassification rate is calculated by summing the off diagonal values and dividing over the number of samples. Please see Appendix B for complete scripts.

Results and Discussion: CART model
Areas of predicted oak wilt are shown by the classification tree below:

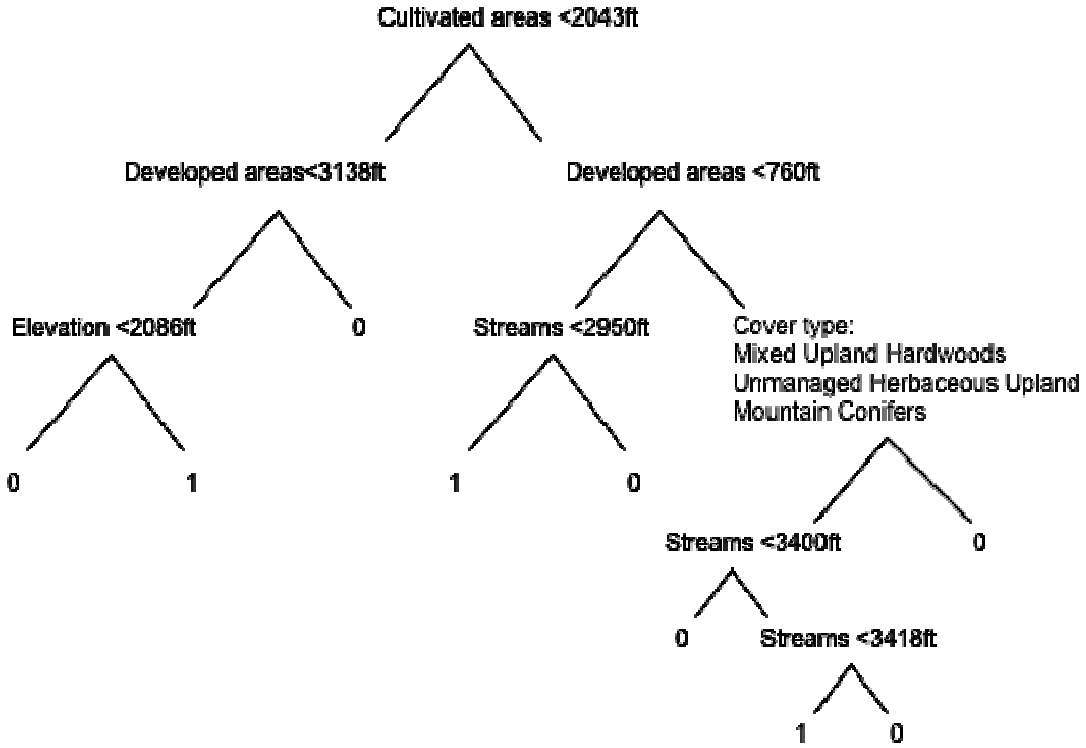


Figure 11 – CART model using "tree" function

Figure 11 displays the branches of a CART model, which are a series of rules that serve to split the sample groups. Everything left of the statement is true, and everything right of the statement is false, and 1 = wilt, 0 = no wilt. If we follow the argument on the left hand side bottom to top, we can see that wilt occurs where elevation is greater than 2,068 ft, distance to developed areas are less than 3,138 ft, and distance to cultivated areas are less than 2,043 ft. Cover type also appeared as a variable with explanatory power, featuring mixed upland hardwoods, unmanaged herbaceous uplands, and surprisingly, mountain conifers. While the mountain conifer cover type suggests that there are no oaks in within this cover type, oaks could be interspersed within these stands and represent their upper elevational distribution limit.

Two statements separated most of the samples; distance from cultivated areas and distance from developed areas (Fig. 11). Elevation, distance to streams, and cover type also served as major classification splits. However, predictors such as TRMI and aspect failed to appear in the model.

I subsequently created a second classification tree using the "rpart" function to see if the predictors that were absent in the first model could potentially appear as splits or alternative splits in this model. The rpart summary report still showed that TRMI, aspect, and slope were not strong predictors of wilt. In addition, the correlation values between TRMI and aspect versus oak wilt are very low (-0.04, 0.05 respectively) while the correlation between distance from cultivated areas and oak wilt is higher in magnitude (-0.4). Therefore, some environmental variables had higher explanatory power than others. Please refer to Appendix B for the full rpart and correlation report. The map below (Fig. 12) displays areas of predicted oak wilt represented by the red shaded areas. Yellow points symbolize the actual oak wilt centers from 1951 to 1997. As seen below, the actual oak wilt points fall within the predicted wilt areas most of the time. Areas for potential expansion of wilt are concentrated in Buncombe County near Asheville, and Jackson County.

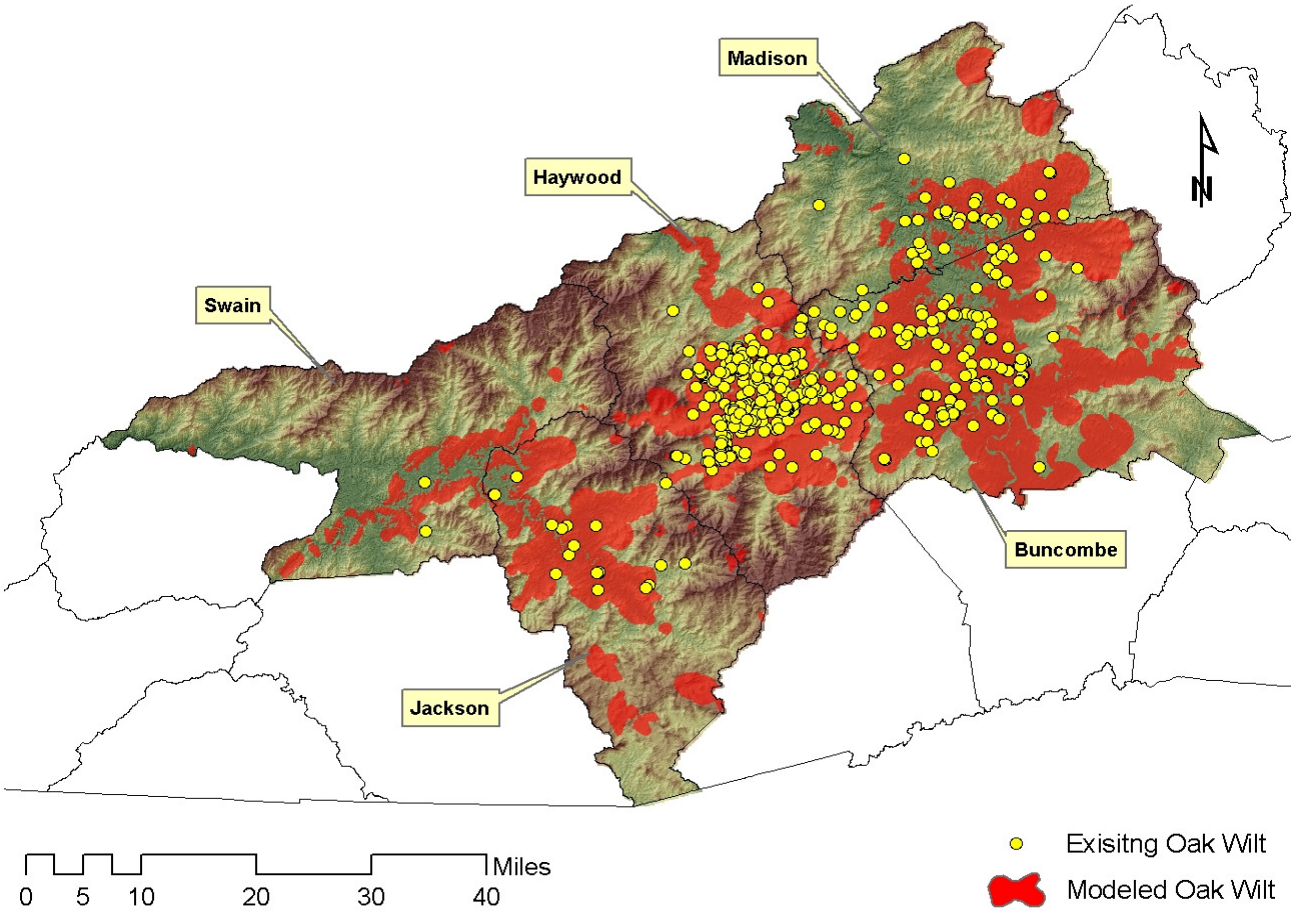


Figure 12 – Map representing predicted areas of wilt shown in red

Given that forests face many threats, such as other diseases and invasive species, this predictive map can help focus efforts on where to concentrate resources for controlling and monitoring oak wilt in the future. Landowners can also easily visualize if their properties are in danger from oak wilt.

Cross-Validation: CART Models

A confusion matrix offers a visual method to see how well the CART model predicted wilt occurrences. The overall success rate of the CART model was 80.9% while the overall misclassification rate was 19.1%. This matrix allows us to see that the model is better at predicting areas of no wilt, rather than areas with wilt. It misclassified 59 samples of no wilt, versus 214 samples of wilt.

Total # Samples: 1430
Success Rate: 80.90%
Misclassification Rate: 19.10%

		Actual	
		No Wilt	**Wilt**
Predicted	**No Wilt**	**530**	**59**
	Wilt	**214**	**627**

Figure 13 – Confusion Matrix for CART model

Objective 3 – Explore which ecological variables are important predictors of oak wilt, using Mantel tests

While the second objective showed that certain environmental variables were useful in predicting where wilt may occur, the model provided no test of statistical significance. In addition, these variables are often correlated with geographic space. For example, while we might find oak wilt in a certain elevation range, is that because certain species are found within that elevation? Furthermore, environmental variables are often intercorrelated among themselves as is the case with TRMI and slope.

To account for these uncertainties, the Mantel (1967) test offers a method to examine the correlation between oak wilt and individual environmental variables, as well as geographic space. It also provides for statistical significance, whereas the CART model did not.

Data Collection and Methods: Mantel Test

Classical statistics are often not adequate for ecological applications because ecology is structured in space. While partial regression can account for correlation among environmental variables, it still fails to address the issue of autocorrelation between environmental variables and geographic space.

Theory of Mantel test: (Legendre and Fortin, 1989)

H_O= Distances among points in a matrix A are not linearly related to the corresponding distances in matrix B.
H_a = Distances among points in matrix A correlated to the corresponding distances in matrix B.

Mantel (1967) Test Statistic:

$$z = \sum_i \sum_j A_{ij} B_{ij}$$

Where A and B represent variables measured at locations i and j (row and column) indices where $i \neq j$.

Normalized Mantel statistic:

$$r = \frac{1}{(n-1)} \sum_i \sum_j \frac{(A_{ij} - \overline{A})}{s_a} * \frac{(B_{ij} - \overline{B})}{s_b}$$

Where s_a and s_b represent standard deviations for variables A and B.

The Mantel test compares two distance or dissimilarity matrices to describe their relationships. Normalizing the Mantel statistic (*r)* allows us to interpret the correlation statistic on a scale from -1 to 1. Highly significant values may also correspond with relatively low correlation values as shown by the Mantel statistic (Dutilleul et al., 2000). Comparing the two matrices allows us to examine the similarity of pairs among sampling stations (Legendre and Fortin, 1989). For this study, the similarity will be whether two sample points are similar in whether they have oak wilt or not, referred from now on as "oak wilt status". There are three main categories of variables in this study, including geographic space or how close together two sample points are, the environment, and oak wilt status. I then applied two variations of the Mantel test to these three variables; the simple and partial Mantel test.

The simple Mantel test asks the following questions:

1. Are two points that are similar in oak wilt status also close together?
2. Are two points that are similar in oak wilt status also environmentally similar?

The partial Mantel test asks the following questions:

1. Are two points that are similar in oak wilt status also close together, after controlling for the environment?
2. Are two points that are similar in oak wilt status also environmentally similar, after controlling for spatial structure?

A dataset of random sample points with no wilt were drawn at random by using the "grid" function in ArcGIS (v.9.1). I then sampled the environmental data using the "sample" function in ArcGIS (v.9.1) to attach environmental data to both sample points showing no wilt, and the actual points featuring wilt. Euclidean distances were calculated using the "dist" function in the Ecodist library (Goslee and Urban, 2006) in R. I initially tried to calculate distance measures using Mahalanobis distances, which takes correlations into account. However, the memory limits in R would not allow me to calculate the distances using this method.

I conducted the simple and partial mantel test with individual environmental variables, and then again with one global environmental variable. To create the global variable, I created a distance matrix for all variables using a rescaled Euclidean distance measurement.

These datasets were then tested using the "mantel" function in R. Please refer to Appendix C for the full script.

Results and Discussion: Mantel Test

Simple Mantel Test – Individual Environmental Variables

Results for the simple mantel test on individual environmental variables are shown below in the path diagram (Fig. 14).

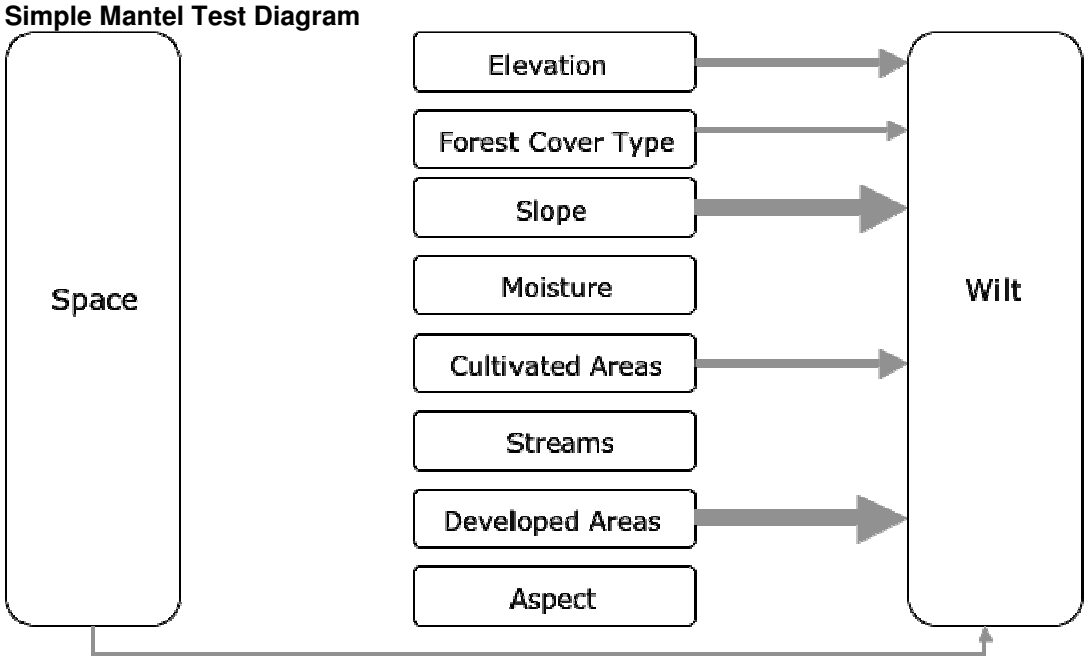


Figure 14 – Path diagram featuring simple mantel test results. Arrows indicate significant relationships (p-value <0.01) and weights indicate degree of correlation between variables.

The simple mantel test asks whether two points that are similar in oak wilt status also close together in geographic space. Figure 14 shows that there is a significant correlation between space and oak wilt status from the bottom arrow linking the two variables ($r = 0.029$). Therefore, two points that are similar in oak wilt status are also close together.

The simple Mantel test also asked whether two points that are similar in oak wilt status also environmentally similar. From Figure 14, we can see that if two points are similar in oak wilt

status, they will also be similar in the following variables: elevation (r =0.071), forest cover type (r = .063), slope (r = 0.104), distance to cultivated areas (r = 0.075), and distance to developed areas (r = 0.102). Magnitudes of correlation were the greatest for distance to developed areas and slope. While distance to developed areas was an expected result from the earlier CART model, showing that two points similar in oak wilt be similar in slope was unexpected. However, it does make intuitive sense that when two points that exhibit wilt may also be close together in space and therefore be on similar slopes.

Overall, the significant Mantel test statistics showed equal to or higher correlation between individual environmental variables with oak wilt status over geographic space.

Partial Mantel Test: Individual Environmental Variables

Isolating the predictor variables through the partial Mantel test clarified some of the autocorrelation found in the simple Mantel tests. Results for the partial Mantel test are represented by the black arrows in the diagram below (Fig. 15).

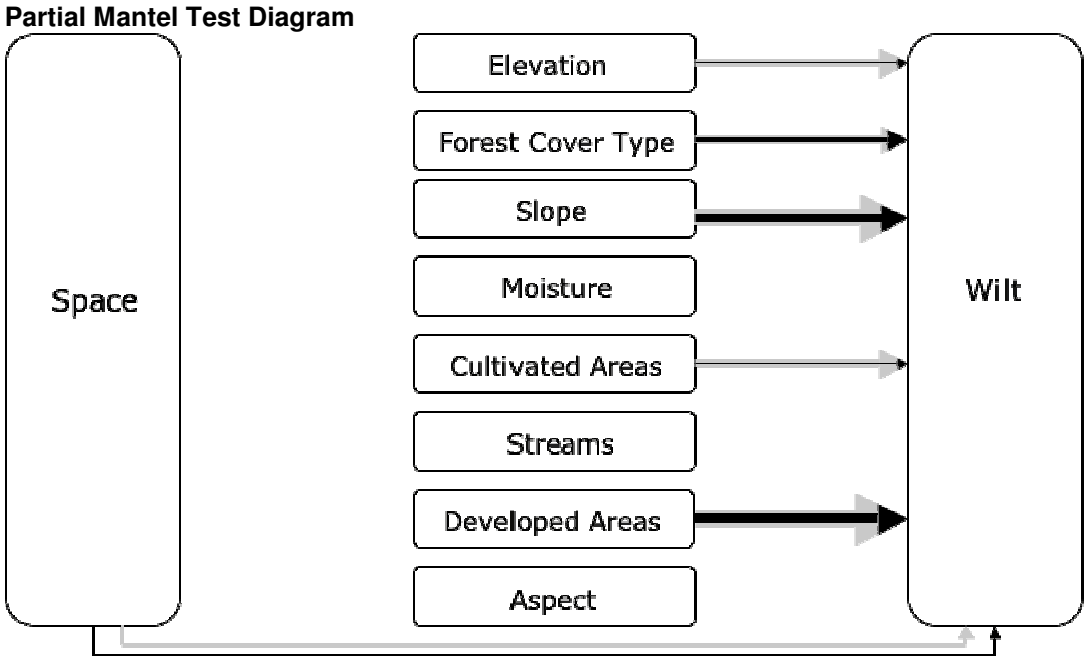


Figure 15– Path diagram featuring partial and simple Mantel test results. Grey arrows represent significant (p-value <0.01) relationships for the simple Mantel test while black arrows represent significant relationships for the partial Mantel test. Weights of the arrows indicate degree of correlation between variables.

Partial Mantel tests allow us to investigate the relationship of two variables while controlling for autocorrelation and intercorrelation between other variables. Figure 15 is similar to Figure 14, but now shows the partial mantel test results in black arrows on top of the grey arrows from the previous simple Mantel test. As we can see, the Mantel test statistic and degree of correlation decreased between all variables.

The first question of the partial Mantel test asked whether two points that are similar in oak wilt status also close together, after controlling for the environment. Similar to the simple Mantel test result, the answer is yes. After taking out the effect of the environment, the closeness of two points in space still affects whether wilt occurs, although the correlation is now only slightly weaker (r =0.0189), as compared to the simple mantel test (r =0.0298).

The second question of the partial Mantel test asked whether two points that are similar in oak wilt status also environmentally similar, after controlling for how close together points are. Results from the partial Mantel test confirm the results of the simple Mantel test. However, Mantel test statistics were lower across all variables after the effect of geographic space was controlled for.

These particular Mantel tests allowed us to test the pure spatial residuals, or the effect of space independent from the environmental variables. While the environmental variables still seem to have a higher impact on whether points are similar in oak wilt status, we can still see the residual effect of space, or how close together two points are. These results also show that environmental variables such as distance to developed and cultivated areas trump geographic space as a predictor variable for wilt, which is somewhat surprising as the disease is locally spread through root graft. From the environmental variables, distance to streams represented the greatest decrease in correlation (-1033%) out of the environmental variables (Fig. 16).

Mantel Test: Oak Wilt Status and Predictor Variables			
Variables	**Simple Mantel Test Statistic (r_s)**	**Partial Mantel Test Statistic (r_p)**	**Percent Change (%)**
Elevation	0.07143005	0.02999778	-58%
TRMI	0.002571083	0.001230153	-52%
Aspect	0.000694505	-0.001055971	-252%
Slope	0.10419901	0.08558466	-18%
Cover	0.06291765	0.06749298	7%
Stream Distance	0.001320685	-0.01232693	-1033%
Developed Distance	0.1023693	0.06561962	-36%
Cultivated Distance	0.07554698	0.1807995	139%
Space	0.02981755	0.01896657	-36%

Figure 16 – Table of simple and partial Mantel test statistics, and percent change.

Simple and partial Mantel test: Global Environmental Variable

All environmental variables were combined to create one global environmental distance measure. By doing so, we can now compare the relationship between oak wilt status, geographic distance with a single environmental variable.

Testing the same questions with the global environmental variable produced similar results as before. Results presented in Figure 17 show that the effect of the environment (r = 0.1068) trumps geographic space (r = 0.0250) (how close two points are to each other) in relation to oak wilt status.

	Oak Wilt	**Environment**	**Space**
Oak Wilt	-	r = 0.1080762 p-value=.001000	r = 0.02981755 p-value = .00100000
Environment	r = 0.10686465 p-value =.001000	-	r = 0.04602101 p-value = .00100000
Space	r = 0.02501665 p-value =.001000	r = 0.04306976 p-value =.00700000	-

Figure 17 – Mantel test with one global environmental variable. The simple Mantel test results are shown in the above diagonal, while partial Mantel test results are shown in the bottom diagonal.

These results confirm previous tests with individual environmental variables, suggesting that some other mechanism in the environment affects the spread of oak wilt more than proximity to infected areas. Therefore, the contagion effect on spread of oak wilt through mechanisms, such as local root grafting, is smaller than specific environmental variables, including development or sap feeding and tree-wounding insects. Given the observations discussed earlier by Liebhold et al. in 1995, we can suggest that soils not conducive to root grafting and the high biodiversity of the Southern Appalachian area contribute to the decreased effect of oak wilt contagion.

CONCLUSION

Oak wilt continues to affect areas of western North Carolina, but three major findings from this project can aid managers in anticipating the spread of wilt in the future. First, oak wilt does prefer specific oak species within the red oak family group. It once preferred *Quercus coccinea* (scarlet oak) up until the 1980s, but now prefers *Quercus rubra* (northern red oak). However, this result depends on the reliability of field data gathered through the Forest Inventory and Analyses (FIA) project and also reflect a change in species composition. Secondly, a predictive model and map were created to show where oak wilt may spread in the future. Buncombe and Jackson country have large areas of potential spread, and managers may want to focus on monitoring and treating these areas if resources are limited for the annual surveys. Lastly, while oak wilt is contagious and spread merely by proximity, environmental variables seem to have a greater impact on whether oak wilt occurs, especially a tree's distance to developed or cultivated areas.

Future studies may want to focus locally on individual oak wilt plots to identify the mechanisms behind preference for northern red oak versus scarlet oak. In addition, studies that focus solely on the relationship between areas disturbed by human impact and oak wilt can help us understand what actions we can avoid to prevent future outbreaks of this disease.

REFERENCES

Appel, D.N. 1995. The Oak Wilt Enigma: Perspectives from the Texas Epidemic. Annual Review Phytopathology. 33:103-18.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and regression trees.* Wadsworth and Brooks/Cole, Monterey, CA.

Boyce, John S. Jr. 1957. Oak wilt spread and damage in the southern Appalachians. Journal of Forestry 55:499-505.

Brown, Mark J. 2002. Forest Statistics for North Carolina. USDA Forest Service Southern Research Station. Forest Inventory Analysis. Resource Bulletin SRS – 88.

De'Ath, G., and K.E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

Degroot, Morris H. 1986. Probability and statistics, second edition. Addison-Wesly Publishing Company, Massachusetts.

Dekker, A. 1978. Spread Patterns of Oak Wilt in Buncombe and Haywood County. Duke University School of Forestry and the Environment.

Dutilleul, P., J. D. Stockwell, D. Frigon, and P. Legendre. 2000. The Mantel Test versus Pearson's Correlation Analysis: Assessment of the Differences for Biological and Environmental Studies. Journal of Agricultural, Biological, and Environmental Statistics. 5:2:131-150.

Engelhard, A.W. 1956. Influence of time of year and type of inoculum on infection of oak trees inoculated with the oak wilt fungus. Plant Dis. Rep. 40:1010-14.

Gibbs, J.N., and D.W. French. 1980. The transmission of oak wilt. U.DA Forest Service Research Paper NC-185, 17 p. USDA Forest Service, North Central Forest Experiment Station, St. Paul, Minnesota. Research.

Goslee, S., and D. Urban. 2006. The Ecodist package. URL: http://cran.r-project.org/doc/packages/ecodist.pdf

Guisan, A., and N.E. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecol. Modeling 135:146-186.

Henry B.W., C.S. Moses, C.A Richards, A. J. Riker. 1944. Oak wilt, its significance, symptoms, and cause. Phytopathology. 34:636-47.

Hepting, George H. 1995. The current status of oak wilt in the United States. Forest Science. 1:2:95-103.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27:209-220.

Legendre, P. and M. Fortin. 1989. Spatial pattern and ecological analysis. Vegetation 80:107-138.

Liebhold, A.M., W.L. Macdonald, D.Bergdahl, and V.C. Mastro. 1995. Invasion by Exotic Forest Pests: A Threat to Forest Ecosystems. Forest Science Monographs 30. 49 p.

R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO) for State [Online WWW] Available URL: "http://soildatamart.nrcs.usda.gov"

Wilson, Dan. 2001. Oak wilt: a potential threat to southern and western forests. Journal of Forestry, 99:5:4-11.

Vayssieres, M.P., R.E. Plant, and B.A. Allen-Diaz. 2000. Classification trees: an alternative nonparametric approach for predicting species distributions. J. Veg. Sci. 11:679-694.

APPENDIX A

Chi-Squared Test Script

1951-1983 versus 1984 FIA data

```
> prop.test(c(6,233,23,167,87), rep(516,5), p=c(26114001, 17512994, 368700, 18923996, 14440001)/77359692)

        5-sample test for given proportions without continuity correction
data:  c(6, 233, 23, 167, 87) out of rep(516, 5), null probabilities c(26114001, 17512994, 368700, 18923996, 14440001)/77359692
X-squared = 585.4493, df = 5, p-value < 2.2e-16
alternative hypothesis: two.sided
null values:
    prop 1      prop 2      prop 3      prop 4      prop 5
0.337565990 0.226383968 0.004766048 0.244623466 0.186660529
sample estimates:
   prop 1     prop 2     prop 3     prop 4     prop 5
0.01162791 0.45155039 0.04457364 0.32364341 0.16860465
```

1984-1997 versus FIA 2002

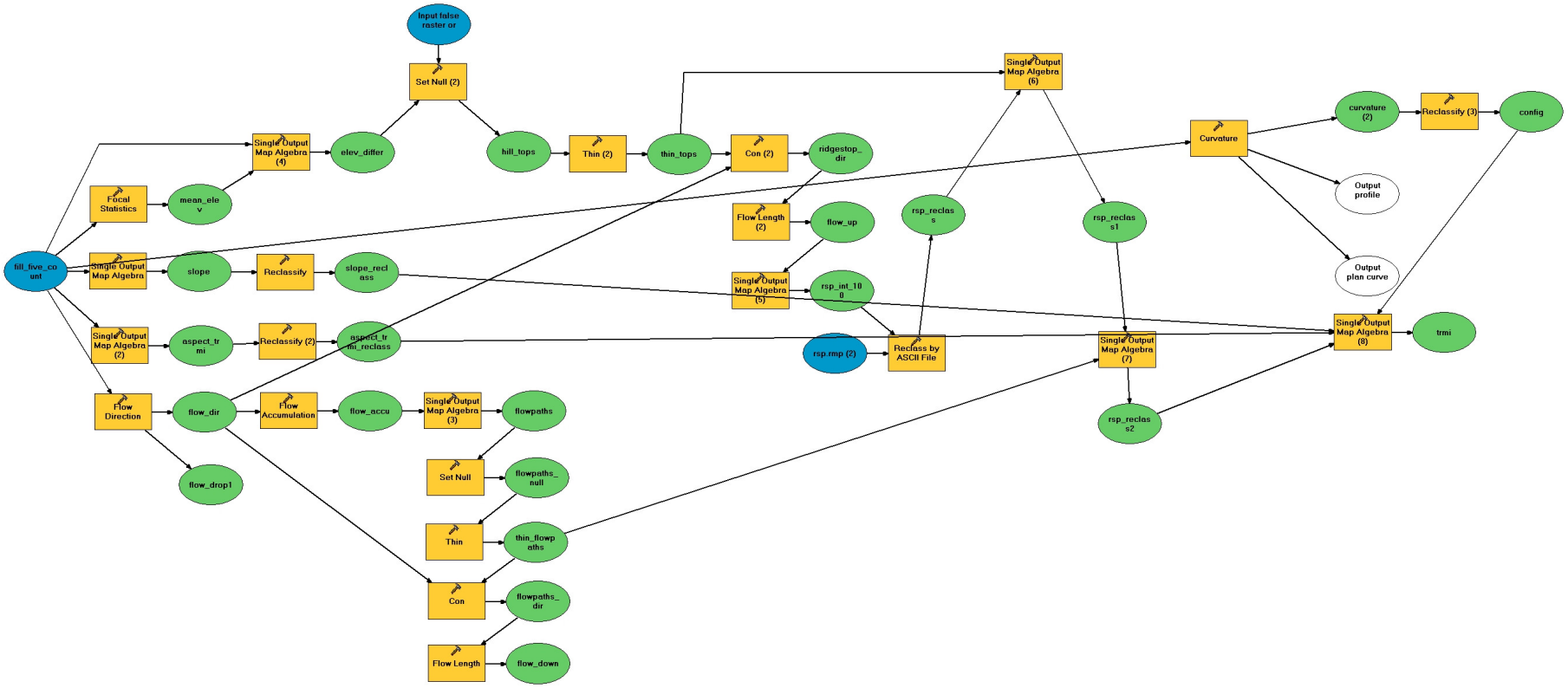
```
> prop.test(c(8,44,9,99,16), rep(176, 5), p=c(6633996,6917991,1032003,10268981,4212001)/29064972)

        5-sample test for given proportions without continuity correction
data:  c(8, 44, 9, 99, 16) out of rep(176, 5), null probabilities c(6633996, 6917991, 1032003, 10268981, 4212001)/29064972
X-squared = 72.6305, df = 5, p-value = 2.903e-14
alternative hypothesis: two.sided
null values:
   prop 1     prop 2     prop 3     prop 4     prop 5
0.22824711 0.23801815 0.03550676 0.35331123 0.14491674
sample estimates:
   prop 1     prop 2     prop 3     prop 4     prop 5
0.04545455 0.25000000 0.05113636 0.56250000 0.09090909
```

Summary Table: Proportion of Oak Wilt Occurence in Scarlet Oak by DBH classes		
DBH Classes (in)	Null population	Oak wilt
5-6.9	0.22969	0.07377
7-8.9	0.22795	0.10655
9.0 - 10.9	0.17996	0.10655
11.0 - 12.9	0.11058	0.05327
13.0 - 14.9	0.07849	0.08196
15.0 - 16.9	0.07921	0.12295
17.0 - 18.9	0.0344	0.21311
19.0 - 20.9	0.01517	0.10655
21.0 - 22.9	0.02269	0.05737
23.0 - 24.9	0.00419	0.04098
25.0 - 26.9	0.00419	0.00081
27.0 - 28.9	0.00925	0.00409
29.0 - 30.9	0.00419	0.02459

APPENDIX B

TRMI Model in ArcGIS (v.9.1)



Python Script: TRMI Model

```
# ---------------------------------------------------------------------------
# TRMI.py
# Created on: Tue Oct 31 2006 08:49:16 PM
#   (generated by ArcGIS/ModelBuilder)
# ---------------------------------------------------------------------------

# Import system modules
import sys, string, os, win32com.client

# Create the Geoprocessor object
gp = win32com.client.Dispatch("esriGeoprocessing.GpDispatch.1")

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")


# Local variables...
slope = "Z:\\MP\\mp_geodatabase.mdb\\slope"
fill_five_count = "fill_five_count"
slope_reclass = "Z:\\MP\\mp_geodatabase.mdb\\slope_reclass"
aspect_trmi = "Z:\\MP\\mp_geodatabase.mdb\\aspect_trmi"
aspect_trmi_reclass = "Z:\\MP\\mp_geodatabase.mdb\\aspect_trmi_reclass"
flow_dir = "Z:\\MP\\mp_geodatabase.mdb\\flow_dir"
flow_drop1 = "Z:\\MP\\mp_geodatabase.mdb\\flow_drop1"
flow_accu = "Z:\\MP\\mp_geodatabase.mdb\\flow_accu"
flowpaths = "Z:\\MP\\mp_geodatabase.mdb\\flowpaths"
flowpaths_null = "Z:\\MP\\mp_geodatabase.mdb\\flowpaths_null"
thin_flowpaths = "Z:\\MP\\mp_geodatabase.mdb\\thin_flowpaths"
flowpaths_dir = "Z:\\MP\\mp_geodatabase.mdb\\flowpaths_dir"
flow_down = "Z:\\MP\\mp_geodatabase.mdb\\flow_down"
mean_elev = "Z:\\MP\\mp_geodatabase.mdb\\mean_elev"
elev_differ = "Z:\\MP\\mp_geodatabase.mdb\\elev_differ"
hill_tops = "Z:\\MP\\mp_geodatabase.mdb\\hill_tops"
```

```
ridgestop_dir = "Z:\\MP\\mp_geodatabase.mdb\\ridgestop_dir"
flow_up = "Z:\\MP\\mp_geodatabase.mdb\\flow_up"
rsp_int_100 = "Z:\\MP\\mp_geodatabase.mdb\\rsp_int_100"
rsp_reclass = "Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass"
rsp_reclass1 = "Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass1"
rsp_reclass2 = "Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass2"
thin_tops = "Z:\\MP\\mp_geodatabase.mdb\\thin_tops"
Input_false_raster_or_constant_value = "1"
rsp_rmp__2_ = "Z:\\MP\\Elevation\\rsp.rmp"
trmi = "Z:\\MP\\mp_geodatabase.mdb\\trmi"
curvature__2_ = "Z:\\MP\\mp_geodatabase.mdb\\curvature"
Output_profile_curve_raster = ""
Output_plan_curve_raster = ""
config = "Z:\\MP\\mp_geodatabase.mdb\\config"

# Process: Flow Direction...
gp.FlowDirection_sa(fill_five_count, flow_dir, "NORMAL", flow_drop1)

# Process: Flow Accumulation...
gp.FlowAccumulation_sa(flow_dir, flow_accu, "")

# Process: Single Output Map Algebra (3)...
gp.SingleOutputMapAlgebra_sa("flow_accu > 25", flowpaths, "Z:\\MP\\mp_geodatabase.mdb\\flow_accu")

# Process: Set Null...
gp.SetNull_sa(flowpaths, flowpaths, flowpaths_null, "VALUE = 0")

# Process: Thin...
gp.Thin_sa(flowpaths_null, thin_flowpaths, "NODATA", "NO_FILTER", "ROUND", "200")

# Process: Con...
gp.Con_sa(thin_flowpaths, flow_dir, flowpaths_dir, "", "VALUE < 1")

# Process: Flow Length...
gp.FlowLength_sa(flowpaths_dir, flow_down, "DOWNSTREAM", "")

# Process: Focal Statistics...
gp.FocalStatistics_sa(fill_five_count, mean_elev, "Rectangle 10 10 CELL", "MEAN", "DATA")
```

```
# Process: Single Output Map Algebra (4)...
gp.SingleOutputMapAlgebra_sa("mean_elev - fill_five_count", elev_differ,
"Z:\\MP\\mp_geodatabase.mdb\\mean_elev;fill_five_count")

# Process: Set Null (2)...
gp.SetNull_sa(elev_differ, Input_false_raster_or_constant_value, hill_tops, "Value > -40")

# Process: Thin (2)...
gp.Thin_sa(hill_tops, thin_tops, "NODATA", "NO_FILTER", "ROUND", "200")

# Process: Con (2)...
gp.Con_sa(thin_tops, flow_dir, ridgestop_dir, "", "VALUE < 1")

# Process: Flow Length (2)...
gp.FlowLength_sa(ridgestop_dir, flow_up, "UPSTREAM", "")

# Process: Single Output Map Algebra (5)...
gp.SingleOutputMapAlgebra_sa("int (( flow_down / ( flow_up + flow_down )) * 100)", rsp_int_100,
"Z:\\MP\\mp_geodatabase.mdb\\flow_up")

# Process: Reclass by ASCII File...
gp.ReclassByASCIIFile_sa(rsp_int_100, rsp_rmp__2_, rsp_reclass, "DATA")

# Process: Single Output Map Algebra (6)...
gp.SingleOutputMapAlgebra_sa("con ( thin_tops == 1, 1, rsp_reclass )", rsp_reclass1,
"Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass;Z:\\MP\\mp_geodatabase.mdb\\thin_tops")

# Process: Single Output Map Algebra (7)...
gp.SingleOutputMapAlgebra_sa("con ( thin_flowpaths == 1, 20, rsp_reclass1 )", rsp_reclass2,
"Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass1;Z:\\MP\\mp_geodatabase.mdb\\thin_flowpaths")

# Process: Single Output Map Algebra (2)...
gp.SingleOutputMapAlgebra_sa("int (aspect (fill_five_count))", aspect_trmi, "fill_five_count")

# Process: Reclassify (2)...
gp.Reclassify_sa(aspect_trmi, "VALUE", "-1 0 0;1 9 18;10 18 19;19 26 20;27 35 19;36 44 18;45 53 17;54 62
16;63 71 15;72 80 14;81 89 13;90 98 12;99 107 11;108 116 10;117 128 9;135 143 7;144 152 6;153 161 5;162 170
```

```
4;171 179 3;180 188 2;189 197 1;198 207 0;208 216 1;217 225 2;226 234 3;235 243 4;244 252 5;253 261 6;262
270 7;271 279 8;280 288 9;289 297 10;298 306 11;307 315 12;316 324 13;325 333 14;334 342 15;343 351 16;352
360 17", aspect_trmi_reclass, "NODATA")

# Process: Single Output Map Algebra...
gp.SingleOutputMapAlgebra_sa("int (slope (fill_five_count))", slope, "fill_five_count")

# Process: Reclassify...
gp.Reclassify_sa(slope, "VALUE", "0 10;3 5.9000000000000004 9;6 8.9000000000000004 8;9 11.9 7;12 14.9 6;15
17.899999999999999 5;18 20.899999999999999 4;21 23.899999999999999 3;24 26.899999999999999 2;27
29.899999999999999 1;30 82 0", slope_reclass, "DATA")

# Process: Curvature...
gp.Curvature_sa(fill_five_count, curvature__2_, "1", Output_profile_curve_raster, Output_plan_curve_raster)

# Process: Reclassify (3)...
gp.Reclassify_sa(curvature__2_, "Value", "-110.75 -5.4033203125 10;-5.4033203125 -2.0322265625 8;-
2.0322265625 1.3388671875 5;1.3388671875 4.7099609375 2;4.7099609375 105 0", config, "DATA")

# Process: Single Output Map Algebra (8)...
gp.SingleOutputMapAlgebra_sa("(aspect_trmi_reclass + slope_reclass + rsp_reclass2 + config)", trmi,
"Z:\\MP\\mp_geodatabase.mdb\\rsp_reclass2;Z:\\MP\\mp_geodatabase.mdb\\aspect_trmi_reclass;Z:\\MP\\mp_geodat
abase.mdb\\slope_reclass;Z:\\MP\\mp_geodatabase.mdb\\config")
```

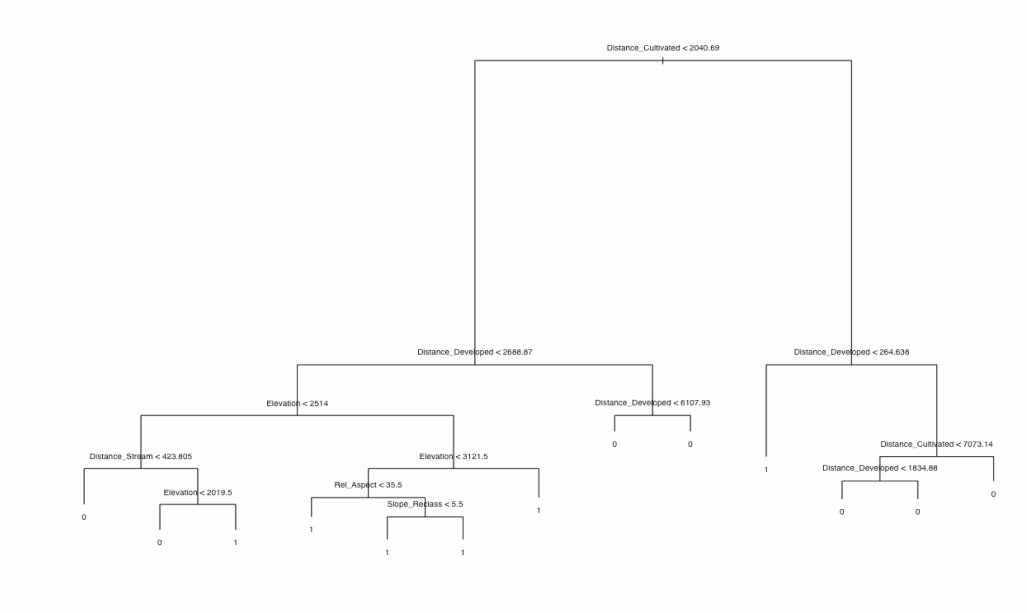
CART Model Script

```
> carttwo=read.csv("carttwo.csv", header=TRUE)
> Occupancy <-carttwo
> occupancy.tree <- tree(as.factor(Occupancy)~., data=carttwo)
> plot(occupancy.tree)
> text(Occupancy.tree, cex=0.8)
Error in text(Occupancy.tree, cex = 0.8) :
        Object "Occupancy.tree" not found
> text(occupancy.tree, cex=0.8)
> summary(occupancy.tree)

Classification tree:
tree(formula = as.factor(Occupancy) ~ ., data = carttwo)
Variables actually used in tree construction:
[1] "Distance_Cultivated" "Distance_Developed"  "Elevation"        "Distance_Stream"     "Rel_Aspect"        "Slope_Reclass"
Number of terminal nodes:  13
Residual mean deviance:  0.7786 = 1112 / 1428
Misclassification error rate: 0.1825 = 263 / 1441
> print(occupancy.tree)
node), split, n, deviance, yval, (yprob)
    * denotes terminal node

 1) root 1441 1994.000 0 ( 0.52394 0.47606 )
   2) Distance_Cultivated < 2040.69 900 1142.000 1 ( 0.33000 0.67000 )
     4) Distance_Developed < 2688.87 781  924.900 1 ( 0.27913 0.72087 )
       8) Elevation < 2514 302  415.300 1 ( 0.44702 0.55298 )
         16) Distance_Stream < 423.805 39   21.150 0 ( 0.92308 0.07692 ) *
         17) Distance_Stream > 423.805 263  348.400 1 ( 0.37643 0.62357 )
           34) Elevation < 2019.5 20    7.941 0 ( 0.95000 0.05000 ) *
```

```
      35) Elevation > 2019.5 243  307.900 1 ( 0.32922 0.67078 ) *
    9) Elevation > 2514 479  441.700 1 ( 0.17328 0.82672 )
    18) Elevation < 3121.5 423  326.900 1 ( 0.13002 0.86998 )
      36) Rel_Aspect < 35.5 49   65.440 1 ( 0.38776 0.61224 ) *
      37) Rel_Aspect > 35.5 374  237.000 1 ( 0.09626 0.90374 )
        74) Slope_Reclass < 5.5 107  116.300 1 ( 0.23364 0.76636 ) *
        75) Slope_Reclass > 5.5 267   91.710 1 ( 0.04120 0.95880 ) *
    19) Elevation > 3121.5 56   77.630 1 ( 0.50000 0.50000 ) *
  5) Distance_Developed > 2688.87 119  151.900 0 ( 0.66387 0.33613 )
  10) Distance_Developed < 6107.93 97  131.500 0 ( 0.58763 0.41237 ) *
  11) Distance_Developed > 6107.93 22    0.000 0 ( 1.00000 0.00000 ) *
3) Distance_Cultivated > 2040.69 541  463.700 0 ( 0.84658 0.15342 )
  6) Distance_Developed < 264.638 57   65.700 1 ( 0.26316 0.73684 ) *
  7) Distance_Developed > 264.638 484  280.800 0 ( 0.91529 0.08471 )
  14) Distance_Cultivated < 7073.14 337  249.500 0 ( 0.87834 0.12166 )
    28) Distance_Developed < 1834.88 100  114.600 0 ( 0.74000 0.26000 ) *
    29) Distance_Developed > 1834.88 237  111.800 0 ( 0.93671 0.06329 ) *
  15) Distance_Cultivated > 7073.14 147    0.000 0 ( 1.00000 0.00000 ) *
```



Docell script to create predictive map in ArcGIS(9.1)

```
&echo &on
docell
if (dem_final_null > 2068.5 & euc_dist_dev < 3138.25 & euc_dist_cul < 2043.96)
pred_wilt = 1
else if (euc_dist_stream < 2950.17 & euc_dist_dev < 760.243 & euc_dist_cul > 2043.96)
pred_wilt = 1
else if (euc_dist_stream < 3418.63 & euc_dist_stream > 3400.15 & lulc = 1,2,10,14 & euc_dist_dev > 760.243 & euc_dist_dev >
2043.96)
pred_wilt = 1
else pred_wilt = 0
end
```

APPENDIX C

Mantel Test Script:

```
library(ecodist)
oak.data<-read.csv("data.csv")
oak.dist<-dist(oak.data[,1])
table(oak.dist)
xy.dist<-dist(oak.data[,3:4])
elev.dist<-distance(oak.data [,4], method="euclidean")
trmi.dist<-distance(oak.data[,5], method="euclidean")
relasp.dist<-distance(oak.data[,6], method="euclidean")
slope.dist<-distance(oak.data[,7], method="euclidean")
cover.dist<-distance(as.factor(oak.data[,8]), method="euclidean")
stream.dist<-distance(oak.data[,9], method="euclidean")
dev.dist<-distance(oak.data[,10], method="euclidean")
cul.dist<-distance(oak.data[,11], method="euclidean")
cover.dist[cover.dist >0] <-1

cor(oak.data)

mantel(oak.dist ~ xy.dist)
mantel(oak.dist ~ elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist + cul.dist + xy.dist, nboot=0)
mantel(oak.dist ~ trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + xy.dist, nboot=0)
mantel(oak.dist ~ relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist + xy.dist, nboot=0)
mantel(oak.dist ~ slope.dist + cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + xy.dist, nboot=0)
mantel(oak.dist ~ cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + xy.dist, nboot=0)
mantel(oak.dist ~ stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + xy.dist, nboot=0)
mantel(oak.dist ~ dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + xy.dist, nboot=0)
mantel(oak.dist ~ cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist + xy.dist, nboot=0)
mantel(oak.dist ~ xy.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist, nboot=0)
```

```
mantel(oak.dist~elev.dist)
mantel(oak.dist~trmi.dist)
mantel(oak.dist~relasp.dist)
mantel(oak.dist~slope.dist)
mantel(oak.dist~cover.dist)
mantel(oak.dist~stream.dist)
mantel(oak.dist~dev.dist)
mantel(oak.dist~cul.dist)
```

Correlation Matrix

	occu	elev	trmi	rel_asp	slope_r	cover	d_s	d_d	d_c
occu	1	-0.30755	-0.04681	0.05845	0.35371	0.26894	-0.02145	-0.44170	-0.40338
elev	-0.30755	1	-0.11072	0.03856	-0.40922	-0.24869	0.35858	0.36360	0.42254
trmi	-0.04681	-0.11072	1	-0.57483	0.23442	0.05708	-0.09562	-0.09056	-0.07305
rel_asp	0.05845	0.03856	-0.57483	1	0.02626	0.00570	0.08254	0.03942	0.01442
slope_r	0.35371	-0.40922	0.23442	0.02626	1	0.21572	-0.21873	-0.36743	-0.33044
cover	0.26894	-0.24869	0.05708	0.00570	0.21572	1	-0.10335	-0.21907	-0.23608
d_s	-0.02145	0.35858	-0.09562	0.08254	-0.21873	-0.10335	1	0.06847	0.07032
d_d	-0.44170	0.36360	-0.09056	0.03942	-0.36743	-0.21907	0.06847	1	0.58297
d_c	-0.40338	0.42254	-0.07305	0.01442	-0.33044	-0.23608	0.07032	0.58297	1

Partial Mantel Tests

```
>mantel(oak.dist ~ xy.dist)
          mantelr   pval1  pval2  pval3   llim.2.5%  ulim.97.5%
       0.02981755   0.001      1  0.001  0.02108539  0.03818862

> mantel(oak.dist ~ elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist +
dev.dist + cul.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
       0.02999778   0.001      1  0.001       0       0

> mantel(oak.dist ~ trmi.dist + relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist +
cul.dist + elev.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
    0.001230153    0.141   0.86  0.279       0       0

> mantel(oak.dist ~ relasp.dist + slope.dist + cover.dist + stream.dist + dev.dist + cul.dist +
elev.dist + trmi.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
   -0.001055971    0.841   0.16  0.331       0       0

> mantel(oak.dist ~ slope.dist + cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist
+ relasp.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
       0.08558466   0.001      1  0.001       0       0

> mantel(oak.dist ~ cover.dist + stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist +
relasp.dist + slope.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
       0.06749298   0.001      1  0.001       0       0

> mantel(oak.dist ~ stream.dist + dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist
+ cover.dist + xy.dist, nboot=0)
          mantelr   pval1  pval2  pval3    llim    ulim
      -0.01232693      1  0.001  0.001       0       0
```

```
> mantel(oak.dist ~ dev.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist
+ stream.dist + xy.dist, nboot=0)
```

mantelr	pval1	pval2	pval3	llim	ulim
0.06561962	0.001	1	0.001	0	0

```
> mantel(oak.dist ~ cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist +
stream.dist + dev.dist + xy.dist, nboot=0)
```

mantelr	pval1	pval2	pval3	llim	ulim
0.01807995	0.001	1	0.001	0	0

```
> mantel(oak.dist ~ xy.dist + cul.dist + elev.dist + trmi.dist + relasp.dist + slope.dist + cover.dist +
stream.dist + dev.dist, nboot=0)
```

mantelr	pval1	pval2	pval3	llim	ulim
0.01896657	0.001	1	0.001	0	0

Simple Mantel Test – Individual Environmental Predictive Variables

mantel(oak.dist~xy.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.02981755	0.001	1	0.001	0.02108539	0.03818862

mantel(oak.dist~elev.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.07143005	0.001	1	0.001	0.06345872	0.08076551

mantel(oak.dist~trmi.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.002571083	0.028	0.973	0.028	0.000922906	0.004268173

mantel(oak.dist~relasp.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.000694505	0.243	0.758	0.543	-0.000945345	0.002615569

mantel(oak.dist~slope.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.10419901	0.001	1	0.001	0.09619996	0.11436281

mantel(oak.dist~cover.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.06291765	0.001	1	0.001	0.0555959	0.07105096

mantel(oak.dist~stream.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.001320685	0.151	0.85	0.268	-0.001194953	0.004061298

mantel(oak.dist~dev.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.1023693	0.001	1	0.001	0.0933842	0.1127025

mantel(oak.dist~cul.dist)

mantelr	pval1	pval2	pval3	llim.2.5%	ulim.97.5%
0.07554698	0.001	1	0.001	0.06613266	0.08458783

Simple Mantel Test: Global Environmental Predictor Variable

```
par.env          <-     mantel(oak.dist~ env.dist + xy.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
     0.10686465   0.001                              1   0.001    0.09823689    0.1158507
par.space        <-     mantel(oak.dist ~ xy.dist + env.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
     0.02501665   0.001                              1   0.001    0.01723072    0.0325885
simp.env         <-     mantel(oak.dist~env.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
      0.1080762   0.001                              1   0.001     0.0994182     0.116824
simp.envspace    <-     mantel(env.dist~xy.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
     0.04602101   0.003                          0.998   0.003    0.03570121    0.0559979
par.spaceenv     <-     mantel(xy.dist ~ env.dist + oak.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
     0.04306976   0.007                          0.994   0.012    0.03238467    0.0528302
simp.wiltspace   <-     mantel(oak.dist~xy.dist)
        mantelr   pval1                          pval2   pval3      llim.2.5%   ulim.97.5%
     0.02981755   0.001                              1   0.001    0.02170096    0.0394127
```