

Using a Geoprocessing Script to Aid in Exploring Correlations between Southern Pine Beetle Outbreaks and Forest Density

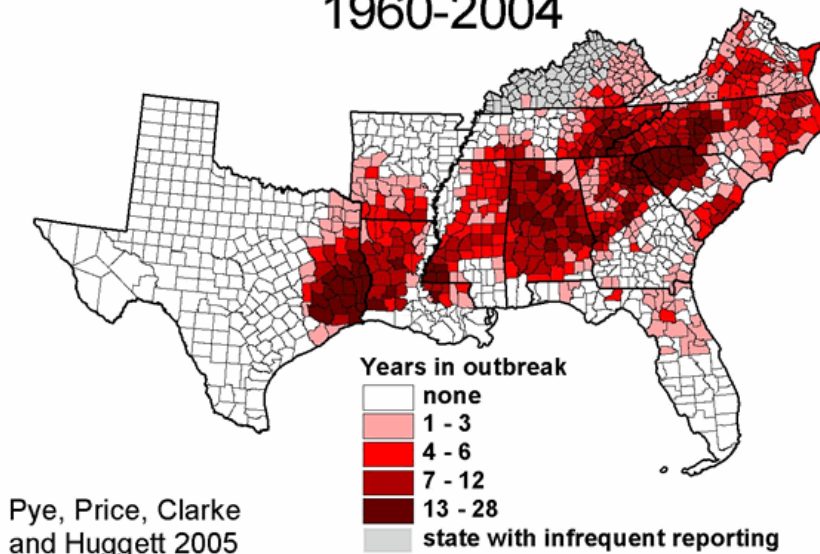
Dahl Winters
December 6, 2006

Introduction

Southern pine beetle (*Dendroctonus frontalis*) is the most destructive insect to pine trees in the southeastern United States. Vast numbers of individuals can become infested and die, with large ecological and economic consequences. Given the prevalence of land fragmentation throughout the southeastern US, it is surprising that the effects of landscape properties on SPB infestations has not been extensively explored. This is an interesting area of research that lies at the intersection of environmental epidemiology and landscape ecology. It is important to know what effects, if any, land fragmentation might be having on SPB infestation severity.

Southern Pine Beetle - Years in Outbreak

1960-2004



If we look at a map of SPB outbreak years by county, counties with the highest number of outbreak years seem to be spatially clustered. This prompted the question of why this clustering was occurring. There had to be some property larger than the county scale but smaller than regional that affected either SPB, host pine populations, or both. An obvious candidate was pine forest density, since the above infestation map seemed to correspond roughly with the distribution of pine forests throughout the southeastern US. However, the distribution of pine forests was known to vary over time, as well as the distribution of the most severe infestations. This prompted further refinement of the initial question.

Problem Statement

How might SPB infestations correlate with the density of various forest types?

Questions that might be answered by answering the above:

- Are SPB infestations more severe in pine plantations or other areas that are predominantly pine forest?
- How might smaller, more mixed stands of pine-hardwood trees contribute to outbreak severity?

Methodology

Data

1. Land Cover Data: This data will be taken from the 2001 National Land Cover Dataset (NLCD 2001), available for free download from the MRLC Consortium.¹ The NLCD 2001 land cover data is derived from 30x30 m resolution Landsat TM imagery. The important land cover classes for this project will be primarily pine forest, which is SPB habitat; classes that contain small or mixed pine stands, which may be potential habitat; and other non-habitat classes such as hardwood forest, agricultural fields, grassland, and development (little forest or cropland to support pest populations). Suitable habitat patches for the maintenance of SPB will be derived from this layer.

2. SPB Infestation Data by County from 1960-2004: This is an Excel spreadsheet (outbrk.xls) that lists counties by year that experienced outbreaks from 1960-2004, providing a temporal and spatial record of outbreaks for the entire southeastern US.²

Approach

Since the goal was to explore how the density of different forest types correlates with SPB outbreak severity, attention will be focused on a single region (NC and SC) and a single year (2001). The year 2001 was chosen for two reasons, despite possible ecological memory effects of prior infestation (areas that had been infested would have dead trees, and would not harbor SPB). First, because the land use-land cover imagery was from 2001, and second, because it identifies at 30 m resolution areas that had live, older trees during that year. These areas indicate places that are SPB-susceptible, since SPB does not infect younger trees. Whether these places have outbreaks has likely to do with the density of SPB arriving at the site from surrounding areas, and the density of trees available to facilitate SPB spread. Other

¹ http://www.mrlc.gov/nlcd_definitions.asp

² <http://www.srs.fs.usda.gov/econ/data/spb/>

factors such as pollution and fire suppression have been ignored for the purposes of this study. The steps to complete this project are as follows:

1) Identify and isolate SPB habitat, potential habitat, and non-habitat from the NLCD 2001 raster dataset. This will be done using reclassification of the NLCD 2001 land cover classes into 3 classes based on the likely availability of pine stands within those classes, according to the NLCD 2001 class descriptions (Appendix A).

2) Identify and isolate SPB infestation data for all counties in NC and SC in 2001. This will be done in Excel first by sorting according to date to isolate all 2001 infestation records. Then, Select Layer by Attributes in ArcMap will be used to isolate all 2001 records for NC and SC ("STATE_NAME" = 'North Carolina' OR "STATE_NAME" = 'South Carolina'), and a spatial join will be done by FIPS code with a NC-SC counties layer. This will provide a map of all 2001 counties that had SPB infestations, and the severity of infestation per area of those counties (using the DISC attribute).

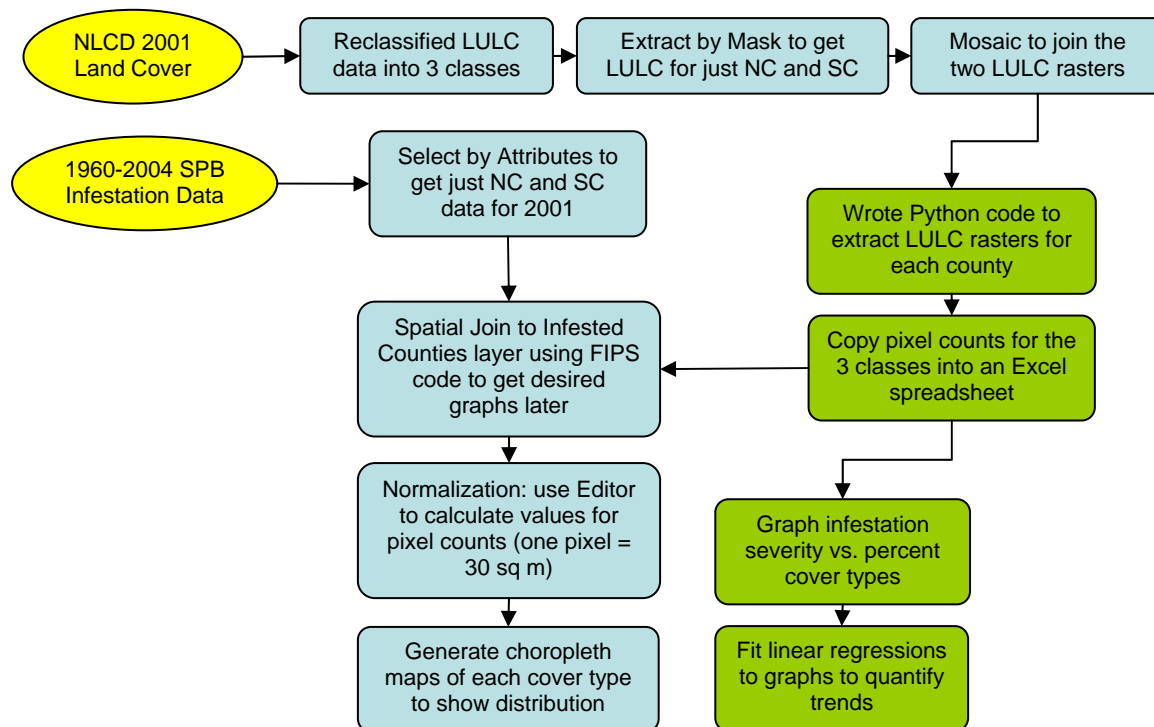
3) Get a count of the pine forest area in each county, and do the same for potential habitat and pine+potential habitat. This will be done by creating individual county shapefiles, then clipping the reclassified land cover raster into individual county rasters using each of the generated county shapefiles. Due to the large number of counties (146) and the large size of the land cover raster, this step would take a very long time and involve many repetitive steps. A batch processing script that would select counties and extract the land cover raster by county would be essential for automating this step.

4) Create county-level maps of the density of pine, potential habitat, and pine+potential habitat. This will be done using the pixel counts from each of the 146 county rasters. Each raster will have 3 classes (1 – pine, 2 – potential habitat, 3 – non-habitat), whose pixel counts will be copied into an Excel spreadsheet. Next, a spatial join will be done to combine the counts of each forest cover class with the counties layer, which contains the infestation severity data.

5) Look to see if outbreak severity correlates with counties that have the most pine forest area per total area. If not, then does outbreak severity correlate with the density of potential habitat, or pine+potential habitat? This will be done by calculating the following 3 percentages using the Editor feature in ArcMap for the counties layer: pine pixels/total pixels, potential pixels/ total pixels, and pine+potential habitat/total pixels. Maps using each of these percentages will be created, using 10-class natural breaks in ArcMap, to visually estimate if there might be any correlations between infestation severity and the density of different forest cover classes. Finally, Excel will be used to graph infestation severity vs. density of forest cover class, and then to apply a linear regression.

It is important to note that this analysis of 2001 land cover and 2001 SPB distributions ignores several things: 1) the effects of ecological memory of past SPB infestations, 2) the contributions of SPB populations in states adjacent to NC and SC, and 3) the effects of any county or state-level pest management efforts in NC, SC, or neighboring states that might reduce outbreak severity in a particular county.

Flow Chart



NLCD 2001 Reclassification

Please see Appendix A for the complete definitions of the land cover types used to create the following three classes:

Class 1. Definitely Habitat for SPB (Predominantly Pine Forest)

Evergreen Forest (42)

Class 2. Potential Habitat for SPB (Small or mixed pine stands)

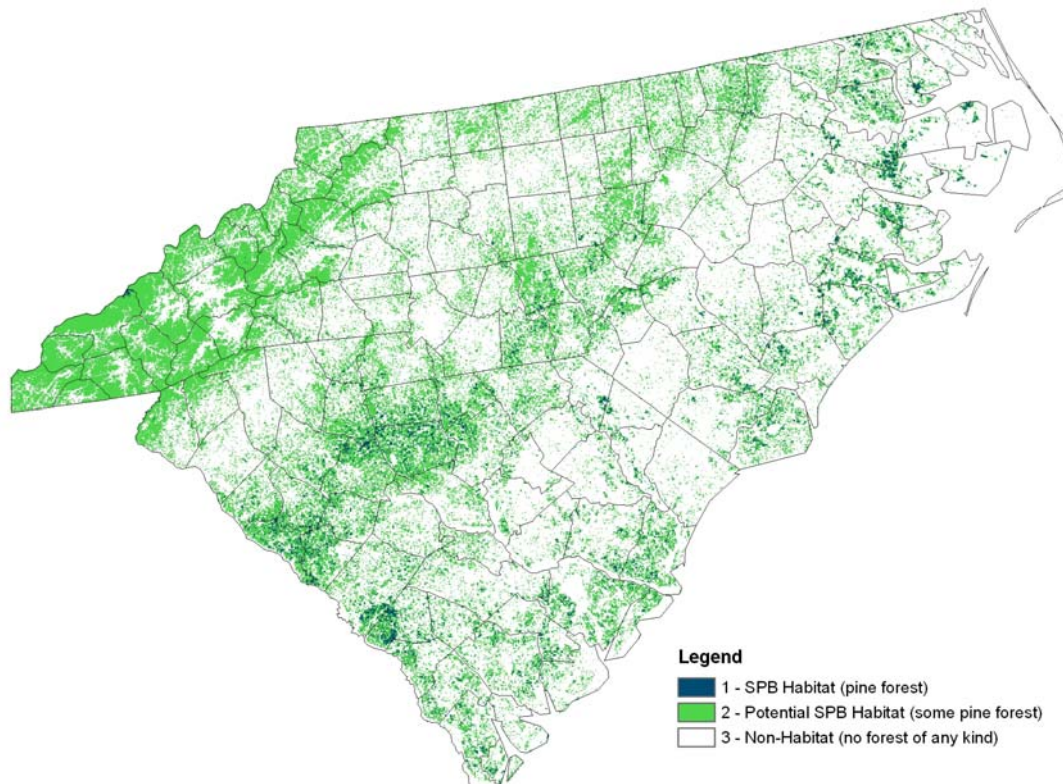
Developed, Low Intensity (22) – SPB often attack stressed and injured trees in older-aged dense stands – some neighborhoods might have these

Deciduous Forest (41) – might be small isolated populations of pine trees here that serve as stepping stones for SPB dispersal
Mixed Forest (43) – likely to have more pine stands than deciduous forest

Class 3. Non-Habitat for SPB:

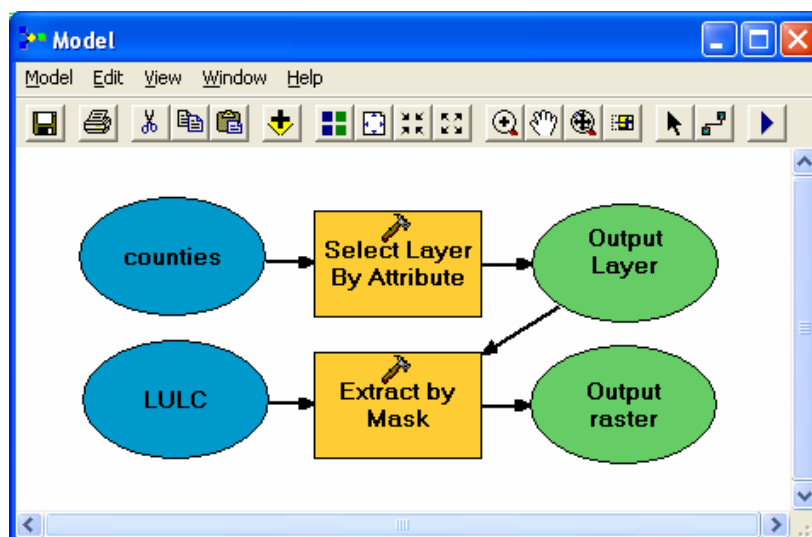
Open Water (11)
Developed, Open Space (21)
Developed, Medium Intensity (23)
Developed, High Intensity (24)
Barren Land (Rock/Sand/Clay) (31)
Shrub/Scrub (52) – since SPB attacks only mature trees, and the trees in this category would be young, early successional ones
Grassland/Herbaceous (71)
Pasture/Hay (81)
Cultivated Crops (82)
Woody Wetlands (90)
Emergent Herbaceous Wetlands (95)

Reclassified Land Cover Raster



Python Geoprocessing Script

This simple Python script for forest extraction originated from ModelBuilder in ArcMap. It was then modified to allow it to loop through each of the 146 counties. Due to heavy processor and virtual memory loads, this script can only process a limited number of counties at a time. Thus, it checks the output folder to see which counties have already been done, before starting the selection and extraction process.



```

# -----
# Forest Extraction Script
# Dahl Winters
# -----

import sys, string, os, win32com.client
gp = win32com.client.Dispatch("esriGeoprocessing.GpDispatch.1")

# Check out Spatial Analyst extension license
gp.CheckOutExtension("spatial")

# Load required toolboxes
gp.AddToolbox("E:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")
gp.AddToolbox("E:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")
gp.AddToolbox("E:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Data Management Tools.tbx")
print "toolboxes loaded..."

# Set workspace
gp.Workspace = "D:/GIS/cntyrasters" #where output files will be stored
print "workspace set..."

# Script arguments...
CountiesLayer = "H:\\GIS\\GIS593\\NCSCcounties-sel-join.shp" #The Infestation Counties
feature layer or any other polygon layer
if CountiesLayer == '#':
    CountiesLayer = "Infestation Counties" # provide a default value if unspecified
print "script args set..."

try:

    # initializing variables
    currentCnty = "Durham"
    currentState = "NC"
    cntyCount = 0
    LULCRaster = "H:\\GIS\\GIS593\\threeclass\\mosaic\\NCSClulc" #The input raster; contains
    #the land cover data to be extracted by county
    print "var init OK, now preparing to start raster extraction..."

    # Create search cursor to go through each county in the Infestation Counties file
    rows = gp.SearchCursor("D:/GIS/infested2.shp")
    #rows = gp.SearchCursor("H:/GIS/GIS593/NCSCcounties-sel-join.shp")
    row = rows.Next()

    while row:
        currentState = row.GetValue("ST") #gets the name of the state, since some counties
        #are in NC and SC
        currentCnty = row.GetValue("CNTYNAME") #gets the name of the county
        print "now working on "+currentCnty, currentState+"...please wait..."

        if gp.Exists(currentCnty+currentState+".shp"):
            print "raster already exists; will skip..."
        else:
            # creates a properly formatted SQL statement for the Select Analysis procedure to
            #follow
            strSQL = "\""+CNTYNAME+"\" + \" = \" + '\"'+currentCnty+'\" + \" AND \" +
            "\""+ST+"\" + \" = \" + '\"'+currentState+'\"

            # selects the individual county from the larger CountiesLayer and saves it as a
            #shapefile for further analysis
            gp.Select_analysis(CountiesLayer, currentCnty+currentState+".shp", strSQL)

            # using the individual county to extract and save the LULCRaster for that county
            # the parameter order is InRaster, InMask, OutRaster
            gp.ExtractByMask_sa(LULCRaster, currentCnty+currentState+".shp",
            currentCnty+currentState)
            gp.BuildPyramids(currentCnty+currentState)
            print currentCnty, currentState+ " complete."

        cntyCount = cntyCount + 1
        row = rows.Next()

```

```

except:
    # If an error occurred while running a tool, then print the messages.
    print gp.GetMessages()

print "The number of counties processed was ", cntyCount
del rows
print "Done"

```

Below is a figure showing the output for a sample shapefile containing just 4 counties:

```

PythonWin 2.1 (#15, Apr 16 2001, 18:25:49) [MSC 32 bit (Intel)] on win32.
Portions Copyright 1994-2001 Mark Hammond (mhammond@skippinet.com.au)
- see 'Help/About PythonWin' for further copyright information.
>>> toolboxes loaded...|
workspace set...
script args set...
var init OK, now preparing to start raster extraction...
now working on Orange NC...please wait...
Orange NC complete.
now working on Durham NC...please wait...
Durham NC complete.
now working on Wake NC...please wait...
Wake NC complete.
now working on Chatham NC...please wait...
Chatham NC complete.
The number of counties processed was 4
Done

```

Maps and Graphs

The following 3 percentages were calculated to generate maps of infestation severity and density of the three forest cover classes:

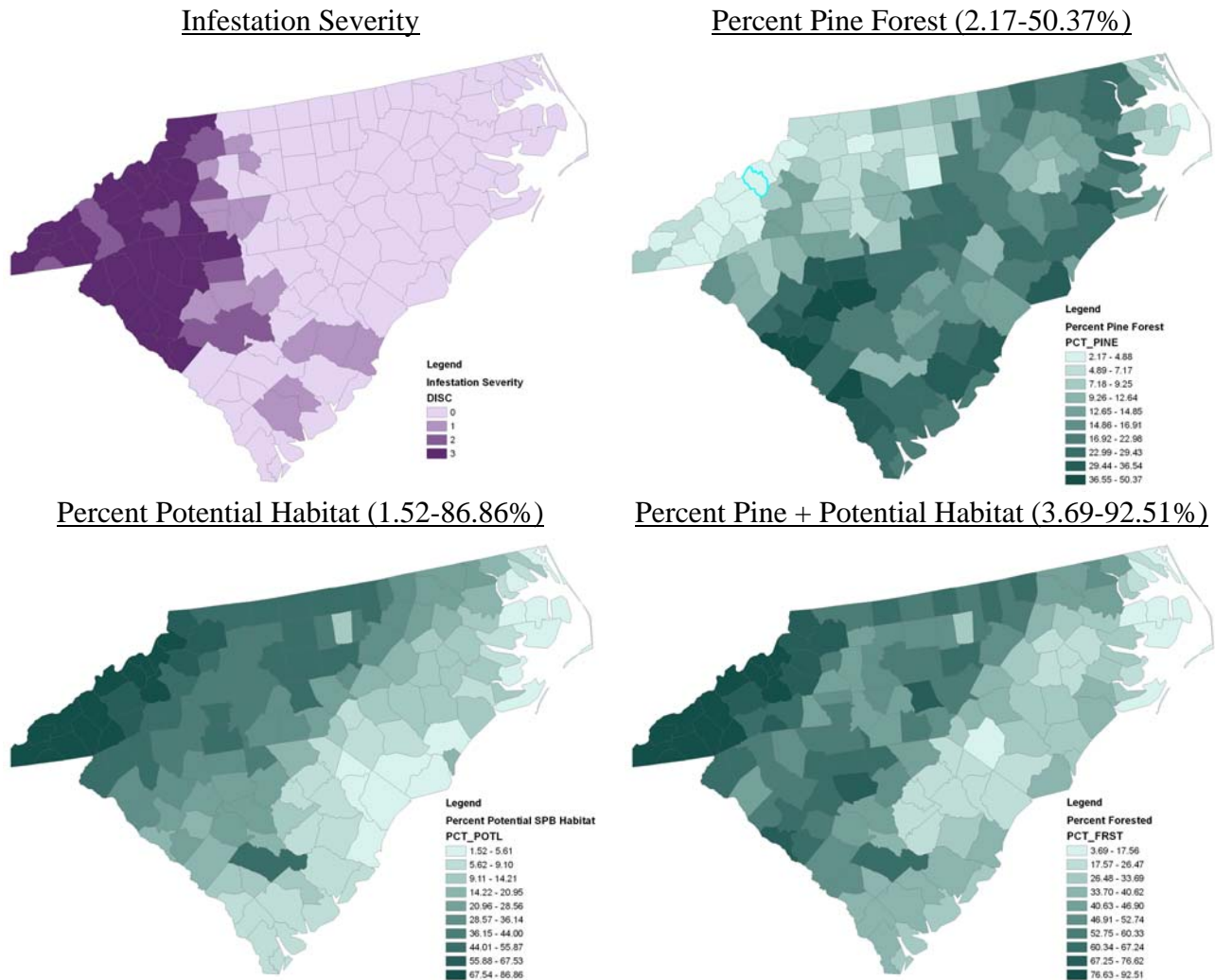
- percent pine forest = pine pixels/total pixels * 100
- percent potential habitat = potential pixels/ total pixels * 100
- percent forested = pine+potential habitat/total pixels * 100

These maps were made to visually estimate if there might be any correlations between infestation severity and the amount of different SPB habitat types.

Results

Comparing Infestation Severity with Forest Cover Class Density

For full-size versions of these maps, please see Appendix B.



A simple visual inspection suggests that generally, infestation severity is inversely correlated with percent pine forest, but directly correlated with percent potential habitat and percent pine + potential habitat. The strongest correlations for all three forest cover types are found in the western NC mountain counties. The northwestern section of South Carolina, which is known to contain large areas of pine forest, does not follow this trend. Instead, infestation severity is more strongly correlated with percent pine forest than the other two cover types. A difference must exist, perhaps in the connectivity of pine vs. potential habitat, between the western NC counties and the northwestern SC counties.

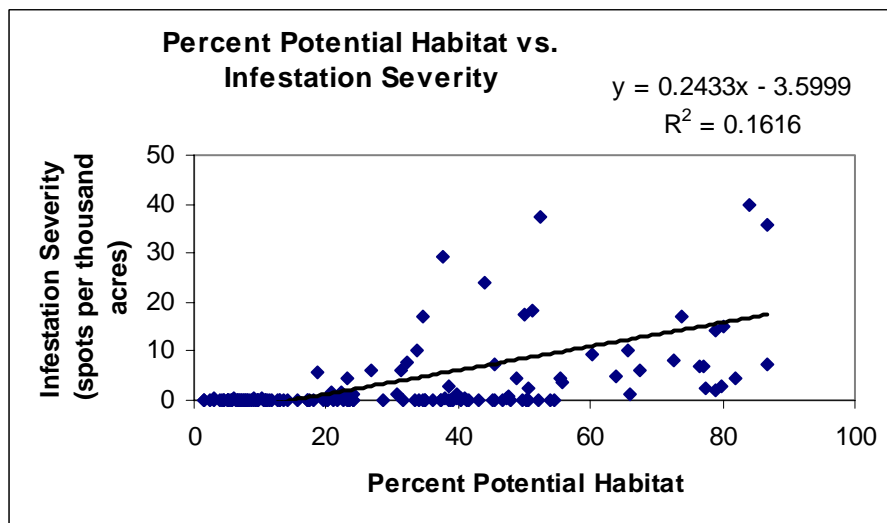
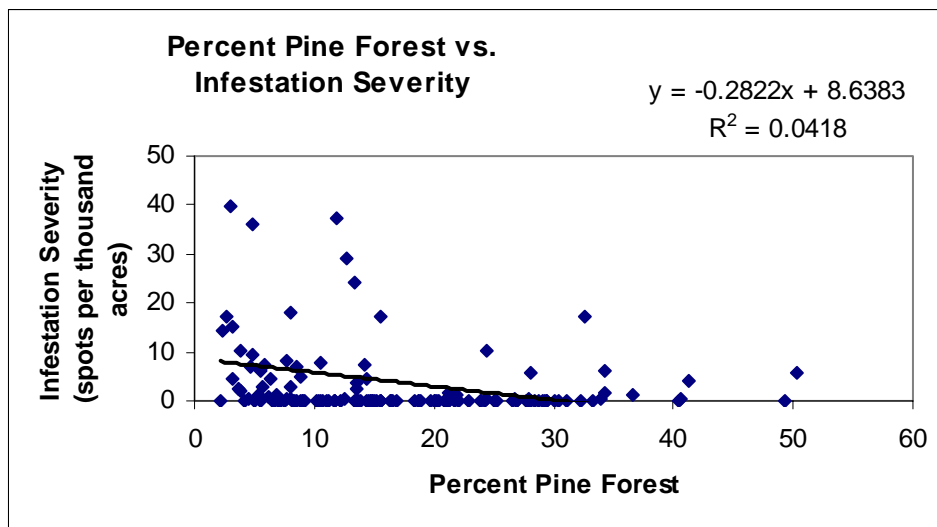
The next Results section describes the trends just described in a more quantitative manner.

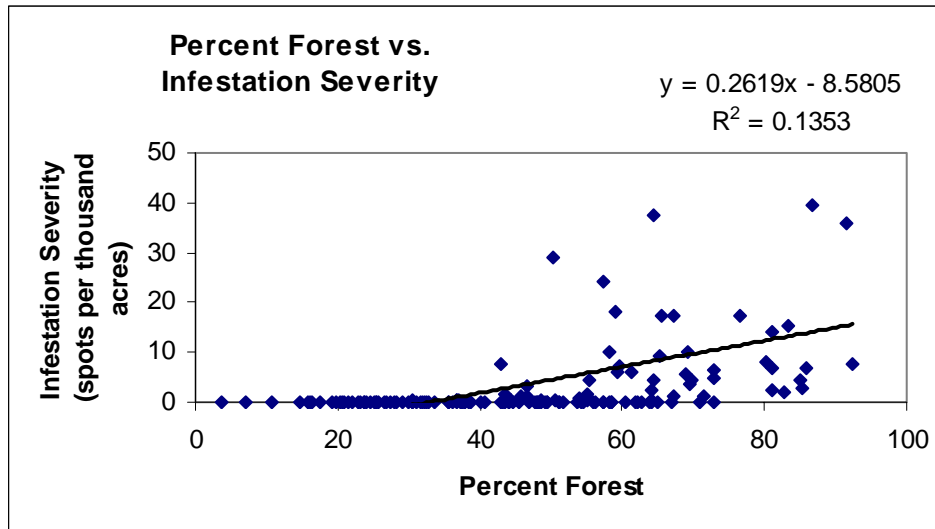
Graphing Infestation Severity Against Forest Cover Class Density

As the percentage of predominantly pine forest increases, interestingly the infestation severity decreases. This is opposite of what would be expected, given that SPB infests pine stands. The greater the density of pine forests in a county, the greater the infestation severity would seem to be.

However, as the percentage of smaller and more mixed pine-hardwood stands increases, infestation severity increases, at a rate of 0.24 more infestations per 1000 acres for every % increase in potential habitat.

As the percentage of both pine forest and smaller, mixed pine-hardwood forest increases, infestation severity increases more strongly, from 0.24 to 0.26 infestations/1000 acres/% increase of forest area.





Discussion

Overall, these graphs suggest that the highest infestation severities occur in counties that have the lowest amount of predominantly pine forest area relative to total area, and instead have a greater number of small or more mixed stands of pine. These small, more mixed stands, which were termed “potential habitat,” may act as SPB population sources and also as stepping stones that facilitate the dispersal of their SPB populations to outbreak sites. A release-capture study by Turchin and Thoeny (1993) found that the maximum flight dispersal distance for 99% of released SPB was 3.29 km.³ Thus, as long as susceptible pine trees exist within 3.29 km of a stand, these trees can serve as population sources. More distant trees can also serve as population sources in rare long-distance dispersal events.

A second theory is offered to explain why infestation severities are highest in counties with lower densities of predominantly pine forest. Perhaps the dominance of small or more mixed pine stands in these counties is a sign of recent land fragmentation. Such fragmentation might create stress on these trees, increasing their susceptibility to SPB. Support for this idea comes from the location of the areas with high infestation severities—many of these counties in western NC and northwestern SC had previously had large, connected tracts of forest that were fragmented by logging and road construction.

How to Interpret and Use Results

These results suggest that prevention of SPB outbreaks should not focus on just pine plantations, or other stands that are dominated by pine. Smaller, more mixed stands of pine

³ Quantifying Dispersal of Southern Pine Beetles with Mark-Recapture Experiments and a Diffusion Model, by Peter Turchin; William T. Thoeny, *Ecological Applications* © 1993 - JSTOR

in both natural and residential areas seem to be more important to SPB outbreaks, perhaps because there are more of them and they are more widely spread.

Sources of Uncertainty

There are two major sources of uncertainty in this analysis. The first comes from pixel misclassification errors in the National Land Cover Dataset. Accuracy levels range from 52% for the bare/mine/transitional land cover class to over 90% for the forest classes. Thus, the greatest accuracies would exist for the pine class, followed by the potential habitat class, and then the non-habitat class. This analysis was concerned primarily with the first two classes. Also, inaccuracies in the land cover data at a 30 meter scale did not seem likely to affect the results of this study, which were produced at a county level.

The second major source of uncertainty comes from the infestation severity data. The generation of this data depended on adequate aerial coverage of each county to spot potential SPB infestations, along with ground-crew follow-ups to confirm SPB infestation at that spot. Infestation severity was calculated as the number of spots per acre, broken up into four severity categories. If some counties are better at spotting SPB infestations than others, this can introduce an unknown amount of uncertainty into this data.

Areas for Further Research

It would be interesting as a followup to this study to trace back through the history of SPB infestations and map the forest cover class densities over time (at least since 1973 when Landsat became available). Then, see whether SPB infestation patterns show the same trends found in this study. The Python script used to extract county forest rasters could be improved upon, to allow users to select the year they want to view infestation severity and forest cover maps for, and to generate a graph of infestation severity vs. forest cover density.

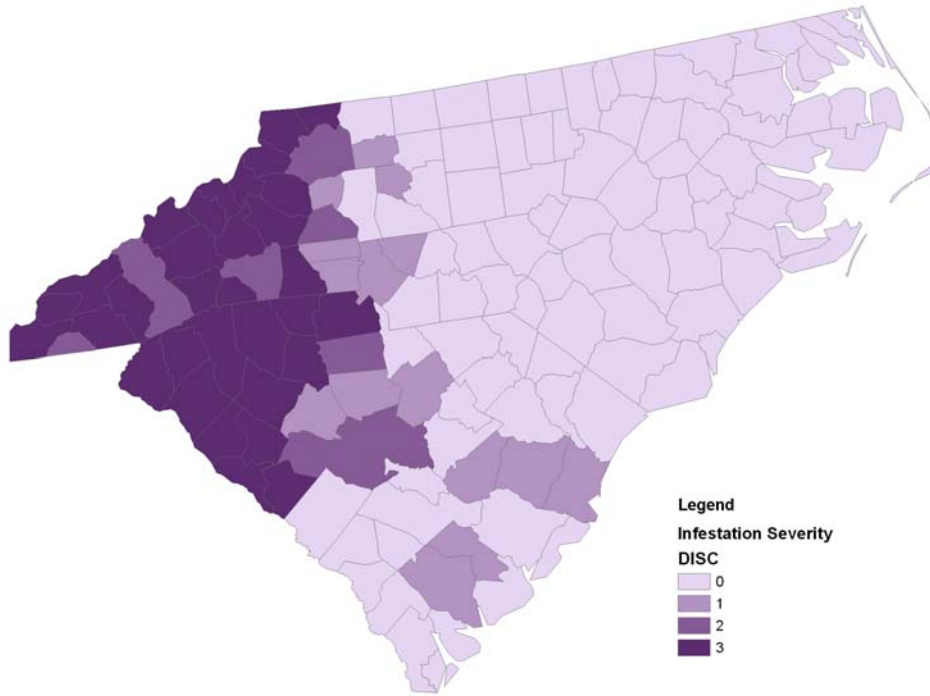
It is unknown exactly how forest density affects SPB outbreak severity. The connectivity of these smaller mixed stands is likely to be more important than we currently think it is, and merits further research into how it might affect SPB spread and outbreak severity. The generated maps identify two regions, western NC and northwestern SC, that might be interesting to look at in terms of pine stand connectivity. If spot-level data were available for counties in these two regions, it would be interesting to use those to see whether there is a threshold distance where SPB spread is greatly enhanced.

Appendix A: NLCD 2001 Land Cover Classes

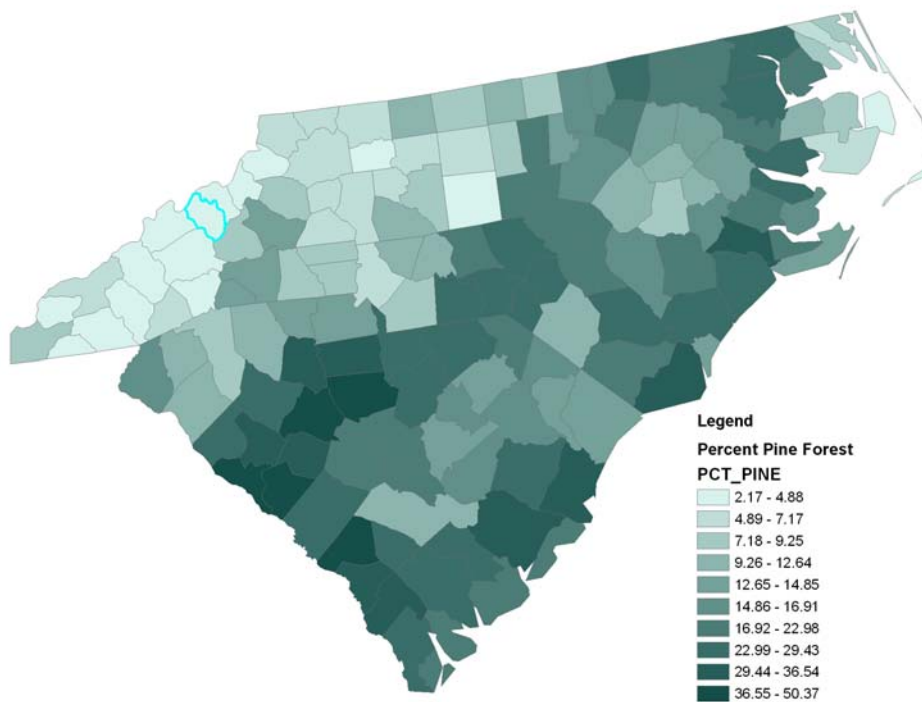
- 11. Open Water** - All areas of open water, generally with less than 25% cover of vegetation or soil.
- 21. Developed, Open Space** - Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes
- 22. Developed, Low Intensity** - Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.
- 23. Developed, Medium Intensity** - Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50-79 percent of the total cover. These areas most commonly include single-family housing units.
- 24. Developed, High Intensity** - Includes highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80 to 100 percent of the total cover.
- 31. Barren Land (Rock/Sand/Clay)** - Barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.
- 41. Deciduous Forest** - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change.
- 42. Evergreen Forest** - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species maintain their leaves all year. Canopy is never without green foliage.
- 43. Mixed Forest** - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75 percent of total tree cover.
- 52. Shrub/Scrub** - Areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.
- 71. Grassland/Herbaceous** - Areas dominated by graminoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.
- 81. Pasture/Hay** - Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation.
- 82. Cultivated Crops** - Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.
- 90. Woody Wetlands** - Areas where forest or shrubland vegetation accounts for greater than 20 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
- 95. Emergent Herbaceous Wetlands** - Areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

Appendix B: Full-Size Maps of Forest Cover Classes

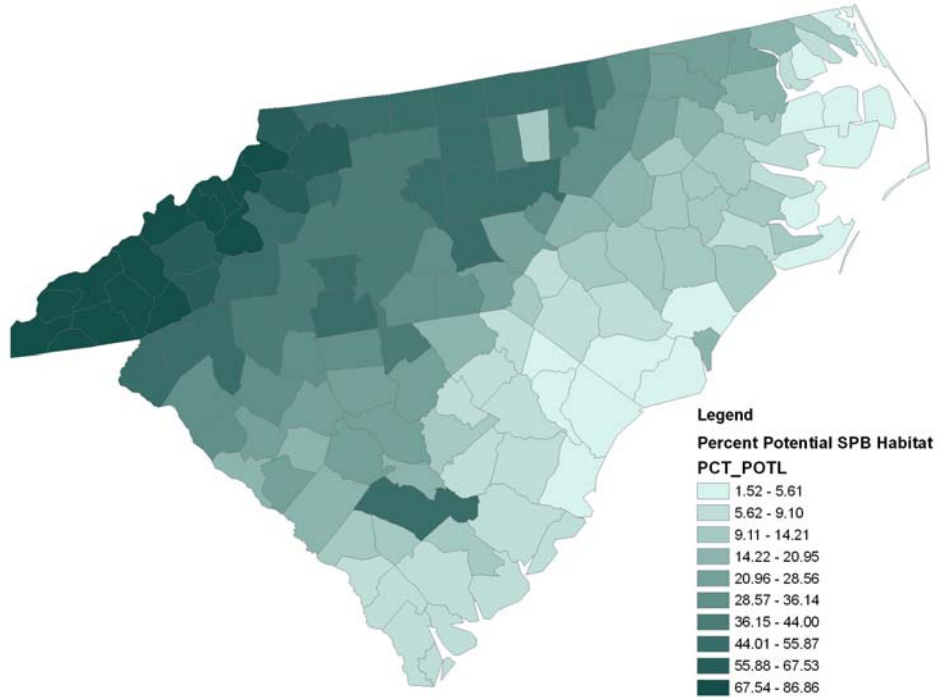
Infestation Severity



Percent Pine Forest



Percent Potential Habitat



Percent Forested (Pine + Potential Habitat)

