Southern pine beetle infestation probability mapping using weights of evidence analysis

Jason Grogan
Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, jgrogan@sfasu.edu

David Kulhavy
Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, dkulhavy@sfasu.edu

James Kroll
Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, jkroll@sfasu.edu

Follow this and additional works at: https://scholarworks.sfasu.edu/forestry

Part of the Forest Sciences Commons

Tell us how this article helped you.

Repository Citation
Grogan, Jason; Kulhavy, David; and Kroll, James, "Southern pine beetle infestation probability mapping using weights of evidence analysis" (2010). Faculty Publications. 228.
https://scholarworks.sfasu.edu/forestry/228

This Conference Proceeding is brought to you for free and open access by the Forestry at SFA ScholarWorks. It has been accepted for inclusion in Faculty Publications by an authorized administrator of SFA ScholarWorks. For more information, please contact cdsscholarworks@sfasu.edu.
INTRODUCTION
The southern pine beetle (SPB) \textit{(Dendroctonus frontalis)} is the most destructive insect pest in the southern forest, (Price and others 1990, Thatcher and others 1980) causing an estimated loss of $265 million in 2001 and $364 million in 2002 (SFIWC 2002, 2003). Historically, SPB populations, and therefore damage, have been high in east TX (Coster and Searcy 1980, Pase 2001).

Predicting where, when and how severe SPB will strike is problematic. Beetle outbreaks are difficult to predict; the best way to reduce loss (hazard) is by determining areas most vulnerable to infestation, then concentrating detection and hazard reduction efforts on the most susceptible areas. Preventing conditions favorable to outbreaks is paramount. Hazard rating models are used to identify forests with characteristics indicative of susceptibility to pests. Hazard maps aid hazard reduction programs by identifying susceptible areas to apply practices for reducing susceptibility.

Numerous systems have been developed to rate stand susceptibility to SPB infestations (Coster and Searcy 1980, Mason and others 1985). Most of these systems are similar; utilizing specific site and stand characteristics to estimate susceptibility to SPB attack. The majority use landform, soil productivity and/or stand density as factors (Mason and others 1985). Major drawbacks are lack of availability of necessary data and poor resolution of hazard maps produced. Past rating systems produced maps with resolutions too poor to be used to identify small individual, yet high-hazard stands, especially those on non-industrial private forest landholdings (NIPF). Many past hazard rating systems have been unable to distinguish between high, moderate and low hazard areas within these “patchworks” of small parcels. Maps often produced generalized hazard ratings reflecting “average” condition among NIPF parcels and stands that were not useful for the small landowner in hazard rating of their individual property. Molnar and others (2003) studied the SPB hazard reduction practices of NIPF landowners, finding one of the key reasons for not performing these practices was lack of knowledge about the problem. Their findings indicate need to identify and educate owners of high-hazard properties. Although past systems were useful for landscape-level hazard rating and identifying specific regions for cultural activities, they have not been useful to the owners of nearly 142 million acres of NIPF land in the Southeastern United States (Wear and Greis 2002). High-resolution (satellite) data, combined with rapid processing ability of today’s geographic information systems (GIS) allow production of hazard maps helpful for even the smallest forest stand.

In 2003, the Research, Development and Applications Agenda for a Southern Pine Beetle Integrated Pest Management Program stated that forest and SPB managers, in general, “…have inadequate knowledge of the usefulness of remote sensing technologies to detect SPB infestations and identify susceptible conditions.” They recommended future research address the following: 1) “…tools to help determine where and which silvicultural protocols should be applied to prevent or reduce SPB-caused impact;” 2) “…more effective methods for monitoring susceptible forest conditions;” and, 3) “SPB hazard and risk assessment protocols improved to enable application at all relevant spatial and temporal scales” (Coulson and others 2003). We incorporate the use of Weights of Evidence spatial analysis (WoF) to address these research recommendations inclusively, in a manner not exhibited by previous systems. According to Coulson and others (1988) traditional hazard rating systems are problematic in they are not tied to a GIS, complicating map production and slowing the process of updating hazard ratings. This system is incorporated directly into a GIS, allowing for timely processing of more complex (and highly predictive) data, rapid hazard updates and efficient map production. WoF produced hazard rating maps, for a larger geographic area, at substantially higher resolution, than are produced by other models.
OBJECTIVES
The main goal was to develop a GIS-based SPB hazard rating system using WofE analysis. The objective was to create a GIS that accurately rates forest stand susceptibility to SPB attack. Data more quickly and efficiently obtained by remote sensing and satellite imagery were used, rather than more time consuming field measurements or photogrammetric interpretation. The final product is a thematic map, predicting probability of southern pine beetle infestation for Angelina, Nacogdoches, San Augustine and Shelby Counties, TX, developed using remotely sensed data. This map could be utilized for SPB prevention and detection by effectively reducing the area in which to concentrate these efforts.

METHODS
The SPB infestation probability model was developed for Angelina, Nacogdoches, San Augustine and Shelby Counties, TX. Total land area is approximately 2.1 million acres (0.85 million ha), of which 1.6 million acres (0.63 million ha) are forested. Forestland ownership is approximately 84 percent private, 16 percent federal, state and local government (USFS FIA 2005). Two national forests (Angelina and Sabine), and two wilderness areas (Turkey Hill and Upland Island) are within the study area.

Weights of Evidence analysis was used to develop SPB occurrence probability maps. WofE has been used extensively for mineral potential mapping, and many other applications, however it had not yet been employed for forest insect hazard prediction. WofE is a data-driven, Bayesian model for spatial analysis, which utilizes multiple input layers and known occurrence locations to calculate the odds of the occurrence in a different geographic or temporal extent. ESRI ArcView® Spatial Data Modeler (Arc-SDM, available for download from http://ntserv.gis.nrcan.gc.ca/sdm/default_e.htm) extension was used to perform the analysis. Resulting probability maps were tested for effectiveness in accurately predicting probability of SPB occurrence.

Acquisition, processing and/or interpretation of numerous existing geographic datasets were required. Data were processed and converted into necessary formats using ESRI ArcMap® (9.0), ArcView® (3.3) and Leica Imagine® (8.7) software. The following GIS data layers, for the years 1992 and 2002, of Angelina, Nacogdoches, San Augustine and Shelby Counties, TX were acquired: forest cover type, forest age, forest patch size, slope percent, SPB occurrences for training data, and a grid of the study area. Forest cover type and age data were derived from Landsat 1–7 MSS, TM and ETM data. Forest patch size was produced by assuming a “clump” of forest the same age and cover type were a “patch.” Percentage slope was derived from a USGS 30-m digital terrain model.

A point dataset for all recorded SPB infestations (10 trees or larger in size) in the study area for 1992 were obtained from the Texas Forest Service (TFS), Forest Pest Management, Lufkin, TX. The year 1992 was chosen for model building because it is the most recent year of substantial SPB activity in the study area corresponding with available forest type and age datasets. Half of the training points were selected randomly for use in training the model; remaining points were used to test the model's effectiveness.

Weights of Evidence Analysis
Step-by-step procedures for WofE analysis, found in Arc-SDM Users’ Guide (Kemp and others 2001) were followed. Weights first were calculated for each data layer. The resulting weights then were used to evaluate usefulness of each data layer and to determine if classes were grouped appropriately. In order for examination of how strongly a theme is associated with SPB infestations, ArcSDM automatically generalized theme weights into a table of contrast values. Contrast values are not used to generate the predicted probabilities, yet are a general indicator of a themes overall positive (+) or negative (-) association with point occurrences.

Once weights were calculated and all themes evaluated, the response theme (Unique Condition Grid) was produced. This grid is a combination of weights from all evidential themes and may be displayed as a map of infestation probability. Posterior probability was calculated as the natural log of the odds of an infestation occurring at random in a cell, modified by the weight calculated for each evidential theme. The response theme (posterior probability) then was symbolized as a map showing the probability of annual southern pine beetle occurrences.

Data Analysis and Map Evaluation
Analyses were performed to test hypotheses concerning the WofE results and to determine if the WofE model could be used successfully to predict Southern Pine Beetle hazard. The output data of interest (posterior probability) is a predicted probability of annual SPB infestation within that pixel. All hypotheses were tested at the alpha = 0.05 level. The 1992 probability map was analyzed and evaluated for effectiveness of accurately predicting probability of SPB occurrence. The additional one-half of SPB occurrence points not used in model development were used as a check for model evaluation.

The first test was for significant differences in predicted SPB occurrence probability between actual occurrence locations and randomly selected points. SAS® Enterprise Guide 3.0 Software (General Linear Model Procedure) was used to perform an analysis of variance (ANOVA), to determine if predicted probability values were significantly greater at actual SPB occurrences than points selected at random.

The next test was conducted to determine if number of SPB occurrences was correlated with predicted probability. Poisson regression analysis originally was chosen because it is most suited for testing for randomness (inverse of correlation) where probabilities of occurrence are small (Zar 1999). SAS® 9.1 software GENMOD procedure was used to perform analyses. Poisson analysis indicated data were not distributed randomly. Lack of randomness may be due to what Zar (1999) referred to as contagious or overdispersed data, indicated by goodness-of-fit values greater than degrees of freedom (98 > 64). Contagious data are clumped, rather than distributed randomly across the landscape. Contagious data may be described by the negative binomial distribution; therefore a negative binomial regression was
was correlated with actual hazard rating based on field correlation analysis to determine if predicted WofE probability Guide® 3.0 software was used to perform Spearman's to the WofE predicted hazard rating. SAS Enterprise (1981). These calculated hazard ratings then were compared plot locations using methods described by Mason and others (collected in early 2003) were used to hazard rate individual evaluated for effectiveness at predicting actual SPB hazard. Forest measurement data from 479 field sample plots (collected in early 2003) were used to hazard rate individual plot locations using methods described by Mason and others (1981). These calculated hazard ratings then were compared to the WofE predicted hazard rating. SAS Enterprise Guide® 3.0 software was used to perform Spearman's correlation analysis to determine if predicted WofE probability was correlated with actual hazard rating based on field measurements.

Satisfactory results were obtained from the 1992 WofE analysis, therefore an up-to-date (2002) SPB probability map was generated by applying the developed model to current (2002) evidential themes. This current probability map was evaluated for effectiveness at predicting actual SPB hazard. Forest measurement data from 479 field sample plots (collected in early 2003) were used to hazard rate individual plot locations using methods described by Mason and others (1981). These calculated hazard ratings then were compared to the WofE predicted hazard rating. SAS Enterprise Guide® 3.0 software was used to perform Spearman's correlation analysis to determine if predicted WofE probability was correlated with actual hazard rating based on field measurements.

RESULTS
WofE analysis calculated a weight for each pixel of each evidential theme. This weight indicated the degree to which the pixel value is associated with training points. Weights were generalized as contrasts values; an average of all weights for a particular theme. Contrasts values indicate the degree to which each theme is associated with the training points. Contrast values for this study, listed in order of most, to least strongly associated, were; forest cover type (3.13), forest patch size (2.94), forest age (2.29) and percent slope (0.55). A contrast value of 3.13 for forest cover type indicates, on average, forest cover type 3.13 times more strongly associated with SPB infestation than would be expected with random probability. Contrast values indicated all evidential themes were associated positively with SPB infestations, with average values greater than expected at random; for all themes, except percent slope. WofE analysis resulted in the calculation of posterior probability (probability of SPB infestation) maps, which are thematic maps with each pixel value indicating the probability of an annual SPB infestation for that pixel (pixel area = 100 m² or 1 ha). Probability values resulting from the 1992 WofE analysis ranged from near 0 to 0.15.

Next, effectiveness of WofE for predicting probability of SPB infestation, by comparing WofE results to actual SPB infestations was tested. The first test for the 1992 data was to determine if WofE predicted probability of SPB infestation was significantly greater for locations where SPB actually occurred versus randomly chosen locations. ANOVA results indicated WofE probabilities were statistically greater at actual infestation locations (P<0.0001). Mean probability values were 6.7 percent for infestations and 3.2 percent for randomly selected locations; predicted probability was over twice as great at actual infestations.

The second test of 1992 data was to determine if there were significantly more SPB infestations in the higher hazard areas than lower hazard areas. Actual SPB infestation density increased with hazard, from 0.022 to 0.101 spots/ km², for the 5-class hazard map and from .030 to .067 spots/ km², for the 3-class hazard map (table 1). For the 5-class hazard map, 20.8 percent of SPB infestations occurred on only 8.9 percent (very high hazard) of the total forested area. SPB infestation density ranged from 45.5 km²/spot for very low hazard areas to 9.9 km²/spot for the very high hazard areas. For the 3-class hazard map, 41.3 percent of SPB infestations occurred only on 26.6 percent (high hazard) of the total forested land area. SPB infestation density ranged from 32.8 km²/spot for low hazard areas to 14.8 km²/spot for the high hazard areas (table 1). Negative binomial regression was used to test for correlation between predicted probability and actual number of infestations (Zar 1999). This yielded a goodness-of-fit value of 57.936 (critical value 81.381) with p=0.6907, indicating the model had a good fit. This analysis also indicated probability was related significantly to number of infestations (p=0.0002).

Finally, current (2002) SPB occurrence probability and hazard maps were produced by applying the model developed for 1992 to current evidential themes. Results of this analysis produced a current (2002) SPB occurrence probability map. These maps also were visualized as 3- and 5-class hazard maps.

A final test was conducted on the 2002 WofE data. Since no current infestation data were available, correlation between predicted SPB hazard and SPB hazard calculated from forest inventory data was tested. Hazard ratings for 479 sample plot locations were calculated using the formula published by Mason and others (1981). The discriminant function values produced by this formula were compared to WofE predicted probabilities. Again, Spearman's test was used. Statistically significant correlations (p<0.0001) were found, with a correlation value 0.67. Although significantly correlated, when compared to Mason and others (1981) hazard rating, classification accuracy was only 34 percent exact agreement for the 5-class system and 55 percent for 3-class system. Direct (exact classification) assessment of accuracy potentially could be misleading. Mason's accuracy was approximately 78 percent and the results of the discriminant function analysis are condensed into classes. These factors considered, a weighted accuracy assessment was performed, which considered similarly classified points as well as points classified exactly the same (example: A point may have been classified low by Mason and very low
by WofE; this is significantly better than if it were classified as high or very high by WofE). This resulted in a classification accuracy of 61 percent for the 5-class and 66 percent for the 3-class WofE hazard maps.

**DISCUSSION**

As expected, contrast values indicated forest cover type was most strongly associated with SPB infestations. Logically, SPB require pine hosts, therefore presence or absence of host trees is of critical importance. Contrast values indicated forest patch size was the second most strongly correlated theme, indicating importance of patch size for SPB, which has not been noted in previous studies.

Probability values, on average, were more than twice as high at locations where SPB infestations had occurred, versus randomly chosen locations. This result indicated WofE predicted probability is substantially higher where SPB occurred.

Correlation analysis of 2002 data indicated as SPB probability increased, calculated (Mason) hazard increased as well. Despite the statistically significant correlation, examination of data revealed some disparity between WofE and Mason hazard ratings. Hazard ratings were similar in the very low and low hazard classes, and slowly diverged as hazard class increased in severity. In general, misclassification occurred predominately as a commission error; that is, most WofE error resulted from over-rating SPB hazard. This over-rating of probability may have been due to conditional dependence among the datasets and the small number of infestations used for training data (Bonham-Carter 1996). However, rarely were points misclassified with WofE ratings lower than Mason's rating. Additionally, it should be noted, although the Mason hazard rating system is the standard rating system used in east Texas, it has a reported accuracy of only 71 percent (Mason 1979). Due to lack of an “exact” standard for comparison, a weighted assessment was performed, which gave consideration, not only to those points classified exactly the same, but also to those (at a lesser extent) whose values were similar. The resulting weighted accuracy was 61 percent for the 5-class and 66 percent for the 3-class hazard maps.

A cursory examination of forest measurements data and WofE evidential themes was conducted to determine if trends, in either data set, existed between misclassified points. Initial concerns were the WofE analysis lacked an estimate of stand density, which historically is important for hazard rating (Hicks and others 1980, Ku and others 1981). Average basal area of the 13 most significantly misclassified (WofE very high class vs. Mason very low class) points was 28 square feet per acre (6.4 m²/ha), which is nearly half as dense as the overall average of 42 square feet per acre (9.6 m²/ha). Initially, this seemed to indicate disparity in classification was due to lack of consideration of stand density in the WofE analysis. Although the misclassification problem may be partially explained by stand density further examination revealed substantial variation in stand density among misclassified points (range from 0 to 144 square feet/acre or 0 to 33.0 m²/ha). This variation could indicate another factor may be confounding the classification. Upon further examination, trends were found in both datasets; misclassified points were consistently in the >18 year age class with average tree heights less than 45 feet (13.7 m), with little variation. This trend seemed to indicate an inadequacy in the forest age dataset used in the WofE analysis. Mason and others (1981) classification of these points was rated very low due mainly to the tree heights less than 50 feet (15.2 m), yet WofE classified them as higher hazard, mainly because they fell into the >18 year age class. The misclassification problem created by the forest age

---

**Table 1**—Area, number of southern pine beetle (SPB) infestations, percentage of area, percentage of infestations and infestations per unit area for each hazard class (3- and 5- class SPB hazard maps) for forested and total areas (1992 Weights of Evidence analysis)

<table>
<thead>
<tr>
<th>Hazard Class</th>
<th>Area (km²)</th>
<th>% Forest Area</th>
<th>% Total Area</th>
<th>SPB Spots</th>
<th>% SPBs</th>
<th>Spots/km²</th>
<th>km²/Spot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Low</td>
<td>1772.9</td>
<td>29.2</td>
<td>49.2</td>
<td>39</td>
<td>14.8</td>
<td>0.022</td>
<td>45.458</td>
</tr>
<tr>
<td>Low</td>
<td>1571.9</td>
<td>25.9</td>
<td>18.5</td>
<td>63</td>
<td>23.9</td>
<td>0.040</td>
<td>24.951</td>
</tr>
<tr>
<td>Moderate</td>
<td>1034.5</td>
<td>17.0</td>
<td>12.2</td>
<td>40</td>
<td>15.2</td>
<td>0.039</td>
<td>25.863</td>
</tr>
<tr>
<td>High</td>
<td>1155.5</td>
<td>19.0</td>
<td>13.6</td>
<td>67</td>
<td>25.4</td>
<td>0.058</td>
<td>17.246</td>
</tr>
<tr>
<td>Very High</td>
<td>542.6</td>
<td>8.9</td>
<td>6.4</td>
<td>55</td>
<td>20.8</td>
<td>0.101</td>
<td>9.865</td>
</tr>
<tr>
<td>Total</td>
<td>6077.4</td>
<td>100</td>
<td>100</td>
<td>264</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>3345.0</td>
<td>55.0</td>
<td>67.8</td>
<td>102</td>
<td>38.6</td>
<td>0.030</td>
<td>32.794</td>
</tr>
<tr>
<td>Moderate</td>
<td>1115.0</td>
<td>18.3</td>
<td>13.2</td>
<td>53</td>
<td>20.1</td>
<td>0.048</td>
<td>21.038</td>
</tr>
<tr>
<td>High</td>
<td>1617.4</td>
<td>26.6</td>
<td>19.1</td>
<td>109</td>
<td>41.3</td>
<td>0.067</td>
<td>14.838</td>
</tr>
<tr>
<td>Total</td>
<td>6077.4</td>
<td>100</td>
<td>100</td>
<td>264</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
dataset could be corrected. Data already are available to add a third age class (middle age class), resulting in three classes: <15 years, 15 - <25 years and ≥25 years, to the 2002 evidential themes. Future inclusion of improved forest age class data will be possible when current SPB infestation data becomes available to train the WofE model; i.e., the utility of the model should improve in the future.

WofE may predict hazard effectively, without producing exactly the same hazard estimates as traditional systems. It is quite possible WofE analysis considers variables and interactions not used in traditional hazard rating (such as forest patch size). Realistically, producing a SPB hazard map at a scale and resolution comparable to WofE, using the Mason system would be virtually logistically impossible. Mason's system, as well as many others, rates each stand separately, using extensive photo interpretation and/or ground-based measurements. Ultimately, the cost and time savings achieved by implementing a WofE system over a traditional system, could easily justify the possibility of over-estimating hazard of some stands. WofE hazard rating’s effectiveness may not be tested truly until SPB infestations again occur in the study area. Although WofE may over-rate hazard in some areas, it effectively reduced the area for concentration of SPB detection and mitigation efforts. Reconnaissance efforts could be reduced to 20 percent of total area and 32 percent of forested area using the 2002 3-class hazard map, and 11 percent and 17 percent, of total and forested area, respectively, using the 5-class hazard map. The resulting cost-savings of using remote sensing to narrow SPB detection/reconnaissance efforts to the most likely infestation areas could be dramatic. WofE hazard rating also should prove to be an effective tool for hazard reduction education and/or cost-share programs, as it can aid in the identification of stands most in need of hazard mitigation.

CONCLUSIONS

WofE effectively, efficiently and rapidly produced high-resolution SPB hazard maps at a fraction of the labor, time and cost of traditional hazard rating systems. Past methods attempting wide-area (landscape level) hazard rating generally were successful, yet either failed to produce hazard maps at spatial resolutions useful to the land manager; or, required extensive, costly ground measurements and/or aerial photography interpretation (Billings and others 1985, Gumpertz and others 2000, Hicks and others 1980, Mason and others 1981, McNulty and others 1998). Weights of Evidence proved to be an effective method of predicting hazard of SPB infestations, at high-resolutions, utilizing existing remotely sensed datasets. This was accomplished without the use of extensive forest inventory or other labor-intensive practices, such as aerial photography interpretation.

Are WofE analysis hazard maps potentially useful for addressing the SPB problem in east Texas? The primary reason for NIPF landowners not implementing actions to prevent and control SPB was their lack of knowledge of the problem (Molnar and others 2003). Also, Cronin and others (1999) reported for control of SPB, area-wide management is necessary for localized control efforts to be effective. These two findings, when considered together, implicate usefulness of WofE analysis, or similar methods. Landscape-level data are needed in order to ensure hazard mitigation practices are applied area-wide; yet individual landowners need to receive information pertinent to their ownership. Despite shortcomings, Weights of Evidence SPB Hazard Rating can effectively address these issues simultaneously; provide broad, landscape-level hazard rating at a resolution useful to even the small NIPF landowner. WofE hazard maps are useful for identification of areas to concentrate mitigation, detection and educational efforts.

Technology is changing at a rapid pace. Many technologies presumed to be out of reach to the average landowner (and researcher) only a few years ago, now are available. Satellite imagery and other forms of remote sensing are more easily available, more affordable and increasing in resolution and quality. Most hazard rating systems were developed long before these data were readily available. As new data and technologies have become increasingly available most SPB researchers have attempted to extract data needed for existing hazard rating systems, rather than developing new methods utilizing the full potential of the data. Perhaps the time has come to explore using these current data and technologies in innovative new systems, rather than attempting to mold data into old methods. Research on utilizing remote sensing and spatial analysis tools, such as WofE, potentially will reduce the need for costly field data collection, while still providing much needed information about the health and condition of natural resources.

LITERATURE CITED


