PATTERNS AMONG THE ASHES: EXPLORING THE RELATIONSHIP BETWEEN LANDSCAPE PATTERN AND THE EMERALD ASH BORER

Susan J. Crocker, Dacia M. Meneguzzo, and Greg C. Liknes

Research Forester (SJC, DMM) and Research Physical Scientist (GCL), U.S. Forest Service, Northern Research Station, 1992 Folwell Ave, St. Paul, MN 55108.
SJC is corresponding author;
to contact, call (651) 649-5136 or email at scrocker@fs.fed.us.

Abstract.—Landscape metrics, including host abundance and population density, were calculated using forest inventory and land cover data to assess the relationship between landscape pattern and the presence or absence of the emerald ash borer (EAB) (*Agrilus planipennis* Fairmaire). The Random Forests classification algorithm in the R statistical environment was used to create a model relating the relative importance of landscape predictor variables to the presence/absence of EAB detected from 2003 to 2009. The dataset was then subdivided, based on quarantine year, to create two subsequent models: (1) data from 2003 to 2007 and (2) data from 2008 to 2009. Model accuracy was 85.3, 91.6, and 89.6 percent, respectively. While population density was ranked as the top predictor variable in all three models, analysis of the models separated by quarantine year showed variation among the other top predictors. Measurements of urban development and forest edge influenced the model more among counties quarantined between 2003 and 2007, and host abundance was an important predictor among counties quarantined in 2008 and 2009.

INTRODUCTION

America's forests are home to more than 8.8 billion ash trees (*Fraxinus* spp.) greater than 1 inch diameter at breast height (U.S. Department of Agriculture [USDA] Forest Service). The distribution of ash is spatially skewed. Ninety-three percent of all ash trees are found in the eastern half of the conterminous United States; 72 percent of these individuals are within the northeastern[1] region. The abundance, functionality, and utility of ash make it an ecologically and economically valuable tree species. As a prominent component of riparian forests, ash plays an important role in reducing runoff and enhancing soil stability (Goforth et al. 2002). Ash is a valuable commercial species used in the manufacturing of numerous wood products, including furniture, pulp and paper, crating, and baseball bats (USDA Animal and Plant Health Inspection Agency [APHIS] 2005). Ash was widely sold as nursery stock and planted in urban settings because of its rapid growth and high tolerance of environmental stress.

[1]Includes Maine, New Hampshire, Vermont, Massachusetts, Connectcut, Rhode Island, New York, Pennsylvania, New Jersey, Delaware, Maryland, West Virginia, Ohio, Indiana, Michigan, Illinois, Wisconsin, Missouri, Iowa, Minnesota, North Dakota, South Dakota, Nebraska, and Kansas.

In recent years, the sustainability of the Nation's ash resource has been threatened by the introduction of a wood-boring beetle identified as the emerald ash borer (EAB) (*Agrilus planipennis* Fairmaire). Native to southeast Asia, EAB was first identified in North America near Detroit, MI, in 2002 (Cappaert et al. 2005). Throughout North America, EAB is a pest of ash. All native ash trees are susceptible to EAB infestation regardless of species, size, or vigor (Poland and McCullough 2006). Tree mortality can be rapid: trees can die within 3 or 4 years of infestation (Kovacs et al. 2009). EAB mortality estimates range in the tens of millions of trees (Poland and McCullough 2006, Kovacs et al. 2009). Thus far, EAB has been identified in 13 states and 2 Canadian provinces.

The spread of EAB in the United States has occurred through artificial and natural means. Long-range dispersal was the result of human transportation of infested firewood and nursery stock, while natural dispersal followed short-range flight of beetles to new hosts (Cappaert et al. 2005). Factors governing natural dispersal include host availability, insect flight capacity, physical barriers, and meteorological conditions (Cappaert et al. 2005).

Of particular interest in this study was the influence of physical obstacles in the dispersal of EAB. The level of forest fragmentation in EAB-infested areas makes it a challenge to understand how EAB populations respond to the arrangement and availability of suitable habitat (Schultz and Crone 2001). Therefore, the purpose of this study was to identify how well landscape characteristics (i.e., landscape pattern metrics) predict the presence or absence of EAB. To accomplish this goal, we used forest inventory data, land cover imagery, and a spatial analysis algorithm to quantify patterns on the landscape. A classification algorithm was then used to identify the relative importance of these potential predictor variables. Our motivation was to assess the role of landscape pattern in the dispersal of EAB and provide insight on habitat suitability and risk.

METHODS

Study Area

To investigate the potential effects of landscape pattern on EAB, we examined county-level data from Illinois, Indiana, Ohio, and Kentucky. Counties ranged from highly fragmented to heavily forested and represented long, short, or no EAB infestation intervals. EAB was detected in Ohio, Indiana, Illinois, and Kentucky in 2003, 2004, 2006, and 2009, respectively. Michigan was not included in the study because EAB is considered “generally” infested throughout all counties in the Lower Peninsula.

Distribution of EAB

Presence/absence of EAB and quarantine establishment was determined for all counties in the study area using state-level data (Illinois Department of Agriculture 2009, Indiana Department of Natural Resources 2009, Kentucky Department of Agriculture 2009, Ohio Department of Agriculture 2009). If EAB was present within a county, the year of detection was also recorded. Since the rules for quarantine establishment varied, the entire county was considered quarantined if any portion of the county (township or a multi-county group) was quarantined by the state. EAB has not been found in all quarantined counties. Counties were classified according to the presence/absence of EAB infestations. Quarantine year was used only as a means of splitting our dataset and was not considered in the classification process.

Landscape Metrics

Landscape metrics are measurements that quantify and describe aspects of landscape pattern (Griffith et al. 2000). In this study, 19 metrics were calculated from various data sources, including raster land-cover data and ground-level plot data (Table 1). Selected metrics describe (1) the amount and spatial arrangement of forest and host availability, and (2) aspects of non-forest landscape pattern, such as population and road density.

Data from the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service were used to calculate county-level estimates of ash trees per acre and ash basal area. Measures of urban development and supplementary measures of forest pattern were gathered from the National Land Cover Database of 2001 (NLCD) (Homer et al. 2004) and datasets by Riemann et al. (2009), where one pixel is equivalent to 30 meters, or approximately 98 feet. Additional data required to calculate airport presence/absence, population density, area of park land, and road density were gathered from the National Atlas of the USA and the Environmental Systems Research Institute (ESRI), Inc., respectively (ESRI 2008, National Atlas of the US 2009). To assess potential changes in host availability by county as EAB dispersal increased over time and space, measures of ash abundance and the percentage of forest land were compared by quarantine year.

Morphological Spatial Pattern Analysis

GUIDOS software (European Commission, Joint Research Centre, Institute for Environmental Sustainability, Ispra VA, Italy) was used to assign forest pixels from the NLCD 2001 to one of seven classes (core, islet, bridge, edge, loop, branch, and perforation). Pixel assignment was completed using morphological spatial pattern analysis (MSPA) algorithms that were based on the connectivity and geometry of the pixels (European Commission 2009, Soille and Vogt 2009). To describe the spatial pattern or degree of fragmentation of forest land, a map of the new pixel classification was produced. Pixel counts of each class were divided by the total number of pixels in each county to obtain percentages of each class by county.

Modeling the Presence/Absence of EAB

A dataset of explanatory variables was constructed from calculated landscape metrics (Table 1). The Random Forests (RF) algorithm in the R statistical software environment (Version 2.8.1 [http://www.r-project.org/]) was used to create a model relating EAB presence/absence to the input variables (Breiman 2001). Cutler et al. (2007) found that RF outperformed commonly used modeling techniques, such as logistic regression, in classifications of lichens, cavity-nesting birds, and the presence of invasive plants. RF provided diagnostics that identified the relative importance of each variable. Refer to Appendix A in Cutler et al. (2007) for a description of the Gini index and its use as an indicator of variable importance. By withholding observations (known as the out-of-bag sample, or OOB) as it builds a suite of classification trees, RF also provided an assessment of the classification accuracy.

Three iterations of the model were constructed using different county combinations in an effort to examine potential differences between these populations. The first iteration included data for all counties in the study area (ALLDATA). After the initial model was run, subsequent model iterations were completed to assess potential changes in county-level attributes. This decision resulted from analysis of mean values of percent forest and ash abundance by county and quarantine year. To accomplish this goal, a second iteration of the model was built using only a portion of the dataset: data for all non-quarantined counties and those counties that were quarantined from 2003 to 2007 (PRE-2008). The final iteration included data for all non-quarantined counties and data from counties quarantined in 2008 or 2009 (POST-2007). Year of quarantine was used to divide the datasets between the PRE-2008 and POST-2007 models; however, it was not used to determine presence or absence of EAB. Presence/absence records were used within the model to determine whether EAB had become established in a county; these records are based on confirmed infestations recorded by each state. The PRE-2008 model was used to forecast the presence/absence of EAB in counties quarantined in 2008 and 2009.

Table 1.—Calculated landscape metrics

Metric	Metric description	Data source
Ash density 1	Ash trees per acre of county land	Forest Inventory and Analysis
Ash density 2	Ash basal area as a percentage of total basal area	Forest Inventory and Analysis
Percent forest	Percent of county area that is forested	NLCD 2001
Percent branch	Percent of forest that is connected at one end to edge, perforation, bridge, or loop	NLCD 2001/ GUIDOS
Percent bridge	Percent of forest that is connected at both ends to different core patches	NLCD 2001/ GUIDOS
Percent core	Percent of forest that is the interior area of a forest patch, excluding forest perimeter	NLCD 2001/GUIDOS

Table 1.—continued

Percent edge	Percent of forest that is the outside perimeter of a forest patch	NLCD 2001/GUIDOS
Percent islet	Percent of forest that is connected and too small to contain core	NLCD 2001/ GUIDOS
Percent loop	Percent of forest that is connected at both ends to the same core patch	NLCD 2001/ GUIDOS
Percent perforation	Percent of forest that is in the inside perimeter of a forest patch	NLCD 2001/ GUIDOS
Percent agriculture	Percent of county area that is classified as agriculture	NLCD 2001 (Riemann et al. 2009)
Percent developed	Percent of county that is classified as urban development	NLCD 2001(Riemann et al. 2009)
Percent forest less than 1 pixel	Percent of forest within 1 pixel of a developed edge (urban development, agriculture, and barren land)	NLCD 2001 (Riemann et al. 2009)
Percent forest 1 to 3 pixels	Percent of forest 1 to 3 pixels from a developed edge (urban development, agriculture, and barren land)	NLCD 2001 (Riemann et al. 2009)
Percent forest greater than 3 pixels	Percent of forest that is greater than 3 pixels from a developed edge (urban development, agriculture, and barren land)	NLCD 2001 (Riemann et al. 2009)
Percent parks	Percent of county area designated as park	ESRI data and maps 2008
Road density	Miles of major roads per square mile	ESRI data and maps 2008
Airports	Presence or absence of an airport in the county	National Atlas of the USA
Population density	Number of people per square mile	National Atlas of the USA

RESULTS

The complete dataset consisted of 402 counties. During the study period (2003-2009), 120 counties were placed under state quarantine. EAB had been positively confirmed in 97 of these counties. Comparison of county-level averages for percent forest and ash density (trees per acre) by quarantine year shows that counties quarantined in 2008 and 2009 had more forest land and a higher number of ash trees per acre than counties quarantined prior to 2008 (Fig. 1).

Results from the ALLDATA model show that overall accuracy was high. Based on the OOB, the RF classifier was able to correctly assign presence/absence to 342 counties (out of a total of 402 counties) for 85.3-percent agreement. As measured by the mean decrease in the Gini index, population density provided the most explanatory power of patterns of EAB presence/absence, followed by percent core, percent forest within 1 to 3 pixels, percent developed, and road density (Table 2).

When the 2008-2009 quarantine counties were excluded from model development (PRE-2008 model), classifier accuracy increased to 91.6 percent. The model performed very well in non-quarantined counties and in highly developed areas. Misclassified counties tended to occur along the boundary between quarantined and

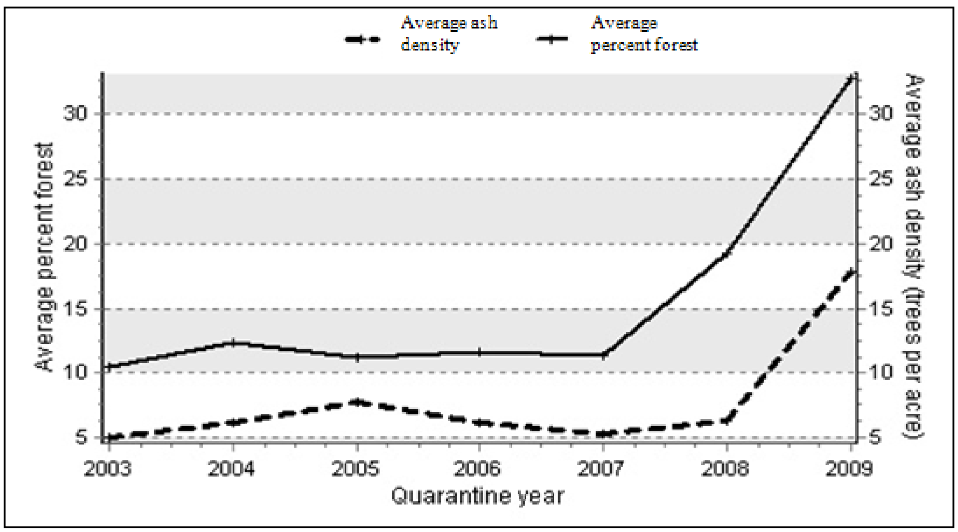

Figure 1.—Comparison of average percent forest and ash density (trees per acre) for quarantined counties by year.

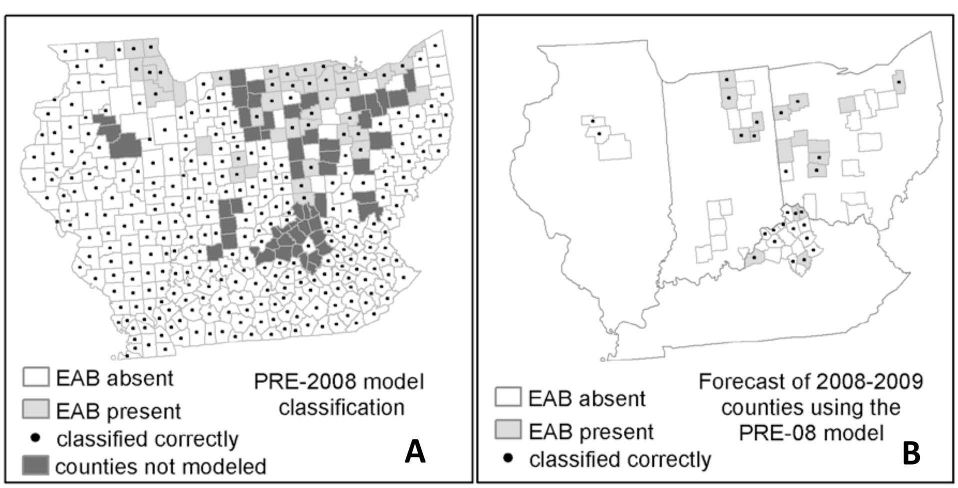

Figure 2.—Classification output from PRE-2008 model (A) and forecast of 2008-2009 counties using the PRE-2008 model (B).

non-quarantined counties (Fig. 2A). Similar to the ALLDATA model, population density was the top-ranked predictor variable; percent developed, percent forest within 1 to 3 pixels, percent core, and percent edge were also highly ranked (Table 2).

The POST-2007 model had a classification accuracy of 89.6 percent (303 out of 338 counties). As with earlier models, population density was identified as the most important explanatory variable with respect to EAB presence. Unlike the previous models, however, measures of host abundance had high predictive value; ash basal area and ash trees per acre ranked third and fourth, respectively (Table 2). Percent forest within 1 to 3 pixels and percent developed were also among the top five predictors, ranking second and fifth, respectively.

When the PRE-2008 model was used to forecast presence/absence in the 2008-2009 quarantine counties, the accuracy was 49.1 percent (27 out of 55 counties). Misclassified counties had no apparent spatial pattern, and classification errors were relatively balanced between false positives and false negatives (Fig. 2B).

DISCUSSION

Overall, models based on landscape metrics performed well in explaining the presence/absence of EAB. Population density, along with percent developed and percent forest within 1 to 3 pixels of a developed edge, were the most important variables in all modeling scenarios, indicating humans play a dominant role in EAB dispersal. Similarly, investigations by Muirhead et al. (2006) found a positive relationship between human population epicenters and the probability of EAB infestation. The PRE-2008 model, which had the highest classification accuracy, also identified percentage of core forest and forest edge as important variables. These findings indicate that forest edges may influence EAB dispersal. Interestingly, investigation of EAB in its native range shows that EAB is known to attack solitary trees and trees along edges (Chinese Academy of Science 1986). Additional research is required, but early analysis of model results indicates that EAB infestations prior to 2008 may be influenced by areas with a high proportion of forest edge.

Conversely, POST-2007 model results show that core and edge metrics, which had high importance in the PRE-2008 model, were replaced by metrics of host abundance as important indicators of EAB presence. The discrepancy between the models highlights a potential change in the relevance of explanatory variables over space and time. Infestations immediately following the initial introduction of EAB may have relied heavily on human activity. Later EAB infestations (those detected after 2007) may have depended heavily on ash abundance. The influence of host abundance is highlighted by an analysis of forestation and ash density, which indicates that the levels of these two attributes remained low and relatively stable for counties quarantined between 2003 and 2007, and then dramatically increased among counties quarantined after 2008 (Fig. 1). Presence of EAB in more heavily forested counties may be related to (1) increased awareness of EAB and a resulting change in human activities or (2) a change in EAB population dynamics in response to more continuously available habitat. Overall, results seem to suggest a change in variable importance over time and/or space and the emerging importance of the level of habitat availability.

This assessment of the relationship between landscape pattern and the presence of EAB has shown that models based on landscape metrics can help provide a measure of the importance of condition-level attributes with regard to EAB presence. However, a change in the explanatory variables across space and time makes it difficult to predict future patterns of infestation. We hope further investigation will yield more consistent results and better predictive models.

Table 2.—Ranking of the importance of explanatory variables (landscape metrics) by model and associated mean decrease in Gini index values.

Order of variable importance (high to low) by model					
ALLDATA model		PRE-2008 model		POST-2007 model	
Variable	Δ Gini	Variable	Δ Gini	Variable	Δ Gini
Population density	21.14	Population density	11.58	Population density	9.41
Percent core	12.11	Percent developed	9.56	Percent forest 1 to 3 pixels	5.89
Percent forest 1 to 3 pixels	12.06	Percent forest 1 to 3 pixels	8.11	Ash density 2	5.50
Percent developed	11.33	Percent core	7.74	Ash density 1	4.82
Road density	10.29	Percent edge	7.51	Percent developed	4.34
Percent edge	9.41	Percent forest greater than 3 pixels	6.14	Road density	4.27
Ash density 2	9.36	Road density	6.07	Percent forest less than 1 pixel	4.23
Percent forest greater than 3 pixels	8.65	Ash density 1	5.76	Percent core	4.17
Ash density 1	8.39	Percent forest less than 1 pixel	4.70	Percent edge	3.83
Percent forest less than 1 pixel	7.53	Percent forest	4.47	Percent forest greater than 3 pixels	3.72
Percent parks	7.52	Ash density 2	4.38	Percent agriculture	3.19
Percent agriculture	6.24	Percent parks	4.34	Percent parks	3.08
Percent forest	6.14	Percent branch	3.66	Percent forest	3.06
Percent branch	4.91	Percent agriculture	3.65	Percent branch	2.51
Percent islet	3.92	Percent islet	2.55	Percent islet	1.77
Percent bridge	2.27	Percent bridge	1.64	Percent perforation	1.12
Percent perforation	1.51	Percent loop	1.32	Percent bridge	1.08
Percent loop	1.48	Percent perforation	1.16	Percent loop	1.05
Airports	0.89	Airports	0.45	Airports	0.48

LITERATURE CITED

Breiman, L. 2001. **Random forests.** Machine Learning. 45: 5-32.

Cappaert, D.; McCullough, D.G.; Poland, T.M.; Siegert, N.W. 2005. **Emerald ash borer in North America: a research and regulatory challenge.** American Entomologist. 51(3): 152-165.

Chinese Academy of Science, Institute of Zoology, 1986. ***Agrilus marcopoli* Obenberger.** Agriculture Insects of China, part I. Beijing, China: China Agriculture Press. 445 p.

Cutler, D. R.; Edwards, Jr., T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. 2007. **Random forests for classification in ecology.** Ecology. 88(11): 2783-2792.

ESRI. 2008. **ESRI® Data and Maps 2008.** Redlands, CA: Environmental Systems Research Institute, Inc.

European Commission, Joint Research Centre. 2009. **GUIDOS.** [online]. Available: http://forest.jrc.ec.europa.eu/download/software/guides.

Goforth, R.; Stagliano, D.; Lee, Y.; Cohen, J.; Penskar, M. 2002. **Biodiversity analysis of selected riparian ecosystems within a fragmented landscape.** Lansing, MI: Michigan Natural Features Inventory. 148 p.

Griffith, J.A.; Martinko, E.A.; Price, K.P. 2000. **Landscape structure analysis of Kansas at three scales.** Landscape and Urban Planning. 52: 45-61.

Homer, C.; Huang, C.; Yang, L.; Wylie, B.; Coan, M. 2004. **Development of a 2001 National Land Cover Database for the United States.** Photogrammetric Engineering and Remote Sensing. 70(7): 829-840.

Illinois Department of Agriculture. 2008. **Fifth amendment to the emerald ash borer interior quarantine**. [online]. Available: http://www.agr.state.il.us/eab/data/200902186640.pdf.

Indiana Department of Natural Resources. 2009. **Emerald ash borer website.** [online]. Available: http:/www.in.gov/dnr/entomolo/3443.htm.

Kentucky Department of Agriculture. 2009. **State quarantines 20 counties to contain emerald ash borer.** [online]. Available: http://www.kyagr.com/pr/newscenter/eabquarantine.htm.

Kovacs, K. F.; Haight, R.G.; McCullough, D.G.; Mercader, R.J.; Siegert, N.W.; Liebhold, A.M. 2009. **Cost of potential emerald ash borer damage in U.S. communities, 2009-2019.** Ecological Economics. doi:10.1016/j.ecolecon.2009.09.004.

Muirhead, J.R.; Leung, B.; van Overdijk, C.; Kelly, D.W.; Nandakumar, K.; Marchant, K.R.; MacIsaac, H.J. 2006. **Modelling local and long-distance dispersal of invasive emerald ash borer *Agrilus planipennis* (Coleoptera) in North America.** Diversity and Distributions. 12: 71-79.

National Atlas of the US. 2009. **National Atlas – Raw data download**. [online]. Available: http://nationalatlas.gov/atlasftp.html.

Ohio Department of Agriculture. 2009. **Ohio's EAB quarantine and regulations.** [online]. Available: http://www.agri.ohio.gov/divs/plant/eab/eab-index.aspx#tog.

Poland, T.M.; McCullough, D.G. 2006. **Emerald ash borer: invasion of the urban forest and the threat to North America's ash resource.** Journal of Forestry. 104(3): 118-124.

Riemann, R.; Lister, T.; Lister, A.; Meneguzzo, D.; Parks, S. 2009. **Development of Issue-relevant Stat level analyses of fragmentation and urbanization.** In: McWilliams, W.; Moisen, G.; Czaplewski, R., comps. Proceedings, Forest Inventory and Analysis (FIA) Symposium 2008; 2008 October 21-23; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

Schultz, C.B.; Crone, E.E. 2001. **Edge-mediated dispersal behavior in a prairie butterfly.** Ecology. 82(7): 1879-1892.

Soille, P.; Vogt, P. 2009. **Morphological segmentation of binary patterns.** Pattern Recognition Letters. 30: 456-459.

U.S. Department of Agriculture, Animal and Plant Health Inspection Agency. 2005. **Emerald ash borer; Quarantine areas.** Federal Register. 70: 249-253.

U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis Program, Forest Inventory and Analysis Data Online (FIDO). [online]. Available: http://fiatools.fs.fed.us/fido/index.html.

U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis Program. [online]. Available: http://www.fia.fs.fed.us/tools-data/default.asp.