DOI: 10.1007/s00267-002-0023-9

Geographic Analysis of Forest Health Indicators Using Spatial Scan Statistics

JOHN W. COULSTON*
Department of Forestry
North Carolina State University
Southern Research Station
Forestry Sciences Laboratory
P.O. Box 12254
Research Triangle Park, North Carolina 27709, USA

KURT H. RIITTERS
USDA Forest Service
Southern Research Station
Forestry Sciences Laboratory
P.O. Box 12254
Research Triangle Park, North Carolina 27709, USA

ABSTRACT / Geographically explicit analysis tools are needed to assess forest health indicators that are measured over large regions. Spatial scan statistics can be used to detect spatial or spatiotemporal clusters of forests representing hotspots of extreme indicator values. This paper demonstrates the approach through analyses of forest fragmentation indicators in the southeastern United States and insect and pathogen indicators in the Pacific Northwest United States. The scan statistic detected four spatial clusters of fragmented forest including a hotspot in the Piedmont and Coastal Plain region. Three recurring clusters of insect and pathogen occurrence were found in the Pacific Northwest. Spatial scan statistics are a powerful new tool that can be used to identify potential forest health problems.

Forest health analysts seek to define the location, extent, and magnitude of changes in forest ecosystems, to explain the observed changes when possible, and to draw attention to the unexplained changes for further investigation. The data come from a variety of sources including satellite images, field plot measurements, and low-altitude aerial surveys. Indicators estimated from the data are assessed in a variety of ways. For example, national assessments typically aggregate information across sites within states, biophysically defined regions (ecoregions), or other large strata. The strata that are characterized by extreme average indicator values become candidates for further investigation.

That approach has some important limitations that will be illustrated for ecoregions. Biophysical parameters often account for some of the normal variation in indicator values, but different types of ecoregions may be better suited for terrestrial or aquatic data, for example, so there is not a single best stratification scheme for all assessment questions (Omernik and Bailey 1997). In addition, health issues that span parts of adjacent ecoregions may be masked by the modifiable area unit problem (Fotheringham and Wong 1991) that dilutes the signal from the affected parts of each ecoregion. Furthermore, the approach may not provide early warning of emerging problems because regional averages will obscure a few extreme values within an ecoregion. Finally, aggregation removes any information about spatial patterns that could have been used to estimate additional indicators for analysis. While regional aggregation schemes are useful for maintaining a common reporting basis over time, it would also be useful to have spatial analysis tools that do not require advance definition of assessment regions. The objective of this study is to describe and demonstrate the use of spatial scan statistics as an analytical tool for identifying spatial clusters with high indicator values.

Epidemiologists and ecologists have long been concerned with detecting and evaluating geographic clusters or hotspots for intensive follow-up analysis. O'Neill and others (1992) applied epidemiological theory to model the spread of disturbances on landscapes, and Flather and other (1998) used geographic overlay techniques to identify regional hotspots containing unusually large numbers of threatened and endangered species. Czaplewski and others (1994) used Moran's **I** statistic to assess spatial autocorrelation of forest growth and detected a spatial cluster of slow-growing forests by partitioning Moran's **I** statistic into its components.

Forest health assessments have much in common with epidemiological studies and several epidemiological analysis methods are potentially applicable to forest health problems. One method is based on a scan statistic to detect clustering in space and time. Developed

KEY WORDS: Monitoring; Spatial analysis; Spatial clusters; Forest fragmentation; Forest insect; Forest disease; Hotspots

Published online May 13, 2003.

*Author to whom correspondence should be addressed at: USDA Forest Service, Forestry Sciences Laboratory, P.O. Box 12254, Research Triangle Park, North Carolina 27709, USA; *email:* jcoulston@fs.fed.us

first for the one-dimensional point process (see Naus 1965), the classical scan statistic examines temporal windows of different length to test the randomness of disease occurrence in time. Kulldorff (1997) extended scan statistics to use geographic windows to permit cluster detection in both the spatial and spatiotemporal domains. Spatial scan statistics have been used to study childhood leukemia (Hjalmars and others 1996), breast cancer (Kulldorff and others 1997) and brain cancer (Kulldorff and others 1998a).

Spatial scan statistics have several advantages for analyzing forest health data over large regions. The primary advantage is that the approach is designed specifically to detect clusters and test their significance; the cluster sizes and regions do not have to be specified in advance. The null and alternative hypotheses are clearly defined, and the test statistic is based on a likelihood ratio and not on an ad hoc procedure. The test is valid regardless of the actual spatial pattern and the approach works with data at multiple spatial scales. The approach is capable of accounting for confounding factors in the background population that are known to be important covariates.

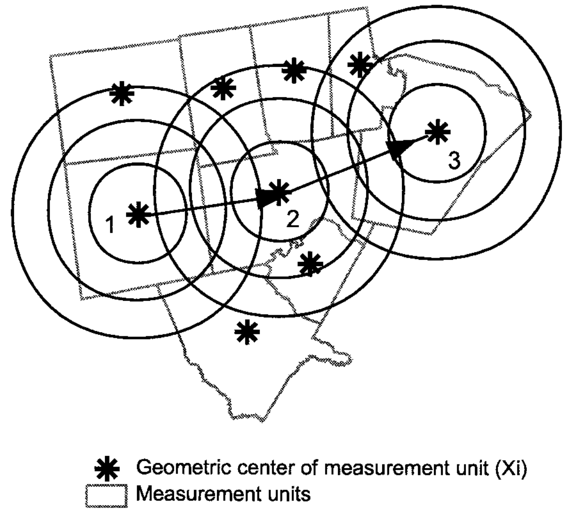

Figure 1. Hypothetical example of the scanning process. In this example, the scanning procedure starts at measurement unit 1 where a set of concentric circles are generated and Ψ_w is calculated for each circle around measurement unit 1. Once the maximum circle size is achieved (say 20% of the area), then the procedure is then carried out at measurement unit 2. This process is carried out for each measurement unit.

Methods

Let i be an index for a set of measurement units that tile a study area, each with a geographic location represented by X_i. Let M_i be the size of a population contained in unit i, and let N_i be the number of individuals in that population that have some attribute of interest. The presence of the attribute of interest for a member of the population will hereafter be referred to as an event. For example, a common application uses counties as the measurement units, and the geographic centroid of each county defines its location. The population of people in each county (M_i) is surveyed to determine how many have the event of a disease (N_i). The objective of the scan statistic is to identify clusters of measurement units for which the occurrence of the event is significantly more likely within the cluster than outside of the cluster.

The scanning proceeds by visiting every X_i (i.e., every location) in the study area. At each X_i, circular windows of different sizes are imposed, with the subject X_i at the center of each one (Figure 1). A window also contains other measurement units if their X_i are within the circle. If there are n_1 measurement units and n_2 windows imposed upon each unit, then the total number of windows in the study area equals $n_1 * n_2$. Each window potentially contains different sets of neighboring measurement units, and each is a potential cluster.

The method for determining the significance of a potential cluster is based on a likelihood ratio. The numbers of events in each measurement unit are assumed to be Poisson distributed (see Kulldorff 1997 for a derivation under the Poisson and Bernoulli distributions). The test statistic for a specific window w (Ψ_w) is defined by the likelihood ratio under the null hypothesis that the rate of events is the same everywhere (Hjalmars and others 1996, Kulldorff 1997, Kulldorff and others 1997):

$$\Psi_w = \left(\frac{\left(\frac{N_c}{M_c}\right)^{N_c} \left(\frac{N_{c'}}{M_{c'}}\right)^{N_{c'}}}{\left(\frac{N^T}{M^T}\right)^{N^T}} \right) I \qquad (1)$$

In equation 1, N and M refer to the number of events and population size, respectively, and the subscripts c and c' refer to the totals of those variables over measurement units within the window, and outside of the window, respectively. $M^T = \Sigma_i M_i$ is the total population size in the study area, and $N^T = \Sigma_i N_i$ is the total number of events in the study area. **I** is an indicator function that has a value of 1 if $N_c/M_c > N_c'/M_c'$, and zero otherwise. The indicator function sets up a one-sided test of the null hypothesis against the alternative that the rate of events is higher within the window.

To understand the test statistic, note that the likelihood function for a specific window is proportional to (Kulldorff and others 1997):

$$\left(\left(\frac{N_c}{\mu}\right)^{N_c}\left(\frac{N^T - N_c}{N^T - \mu}\right)^{(N^T - N_c)}\right)I \qquad (2)$$

where μ is the expected number of events within the window under the null hypothesis that the rate of events is constant across the study area, and **I** is an indicator function that in equation 2 has a value of one if $N_c > \mu$ and zero otherwise. N_c/μ and $(N^T - N_c)/(N^T - \mu)$ are proportional to the event ratios within and outside the window, respectively. For fixed N^T and μ, the likelihood increases with the number of events in the window (N_c).

The window corresponding to the maximum likelihood ratio (i.e., the maximum Ψ_w for all w) identifies by its component measurement units the most likely or primary cluster. Secondary clusters are identified by the ranks of the Ψ_w for all w. The indicator function guarantees that the cluster rates are higher (not lower) than expected under the null hypothesis. The distribution of the maximum Ψ_w and a simulated P value are obtained by a Monte Carlo simulation that repeats the analysis for a large number of random replications of the original data set under the null hypothesis of complete spatial randomness (Kulldorff 1997). The significance test for the primary cluster compares Ψ_w for the primary cluster to the distribution of Ψ from the Monte Carlo simulation. If the value of Ψ_w exceeds 95% of the values from the Monte Carlo simulation, then the cluster is considered significant at the 5% level. Simulated P values for secondary clusters are also obtained by comparing their Ψ_w to the same simulated distribution; those values are considered approximate and conservative estimates (Kulldorff 1997).

For interpreting the results, it is helpful to examine a map of the relative risk for the individual measurement units. Relative risk (R_i) is defined as the ratio of the observed to expected number of events under the null hypothesis (equation 3):

$$R_i = \frac{N_i}{rM_i} \qquad (3)$$

where N_i and M_i are as previously defined, and $r = N^T/M^T$ is the estimated rate of event occurrence over the entire study area. The scanning procedure detects spatial clusters that are local neighborhoods containing measurement units with high relative risk values.

To extend the procedure to three dimensions, the scanning uses cylinders rather than circular windows. Imagine a stack of maps where each layer in the stack represents a different time. The base of the scanning cylinder represents geographic space and the height of the cylinder represents time. The scanning procedure allows both the base and the height to vary continuously as scanning progresses through space and time. There is no change in either the likelihood ratio or the significance test.

The scan statistic approach is computationally intensive, and as a result, several features of the scan statistic are implementation-dependent. This study used the SaTScan software (Kulldorff and others 1998b, Glaz and Balakrishnan 1999) that permits the user to constrain the maximum circle size. One convention is that a circle includes no more than half of the total population (Kulldorff and others 1997). In principle, the size of the circle may vary continuously, but in practice, the test statistic only needs to be computed for circles that contain different subsets of measurement units. The software does not report secondary clusters that overlap the most likely cluster because they provide little additional information. Depending on the data and problem structure, the software indicates the location of a cluster but usually not its exact boundaries.

The scan statistic method is superficially similar to a fractal analysis of mass-area scaling (Milne 1992) and to lacunarity analysis (Plotnick and others 1993) in that all of these procedures involve the counting of events within different-sized windows imposed everywhere in a study area. Fractal analysis is most often used to quantify if and how the rate of events scales in relation to window size, and lacunarity analysis is designed to describe the clumpiness of events over the entire map. Where both of those techniques assume second-order stationarity, the scan statistic is specifically designed to detect clusters that represent nonstationarity.

Applications

We analyzed two aspects of forest health using spatial scan statistics. The first example demonstrates a purely spatial analysis of forest fragmentation patterns in the southeastern United States and the second example demonstrates a space–time analysis of insect and pathogen disturbance over a 10-year period in the Pacific Northwestern United States (Figure 2). These indicators were selected because they are part of the forest health assessment framework specified by the Montréal Process and, as such, are used for strategic forest planning (USDA Forest Service 2000) and national assessments of forest sustainability (e.g., USDA Forest Service 2003), resource conditions (e.g., USDA Forest Service 2001), and forest health (e.g., Conkling and others 2003). According to this framework, forest fragmenta-

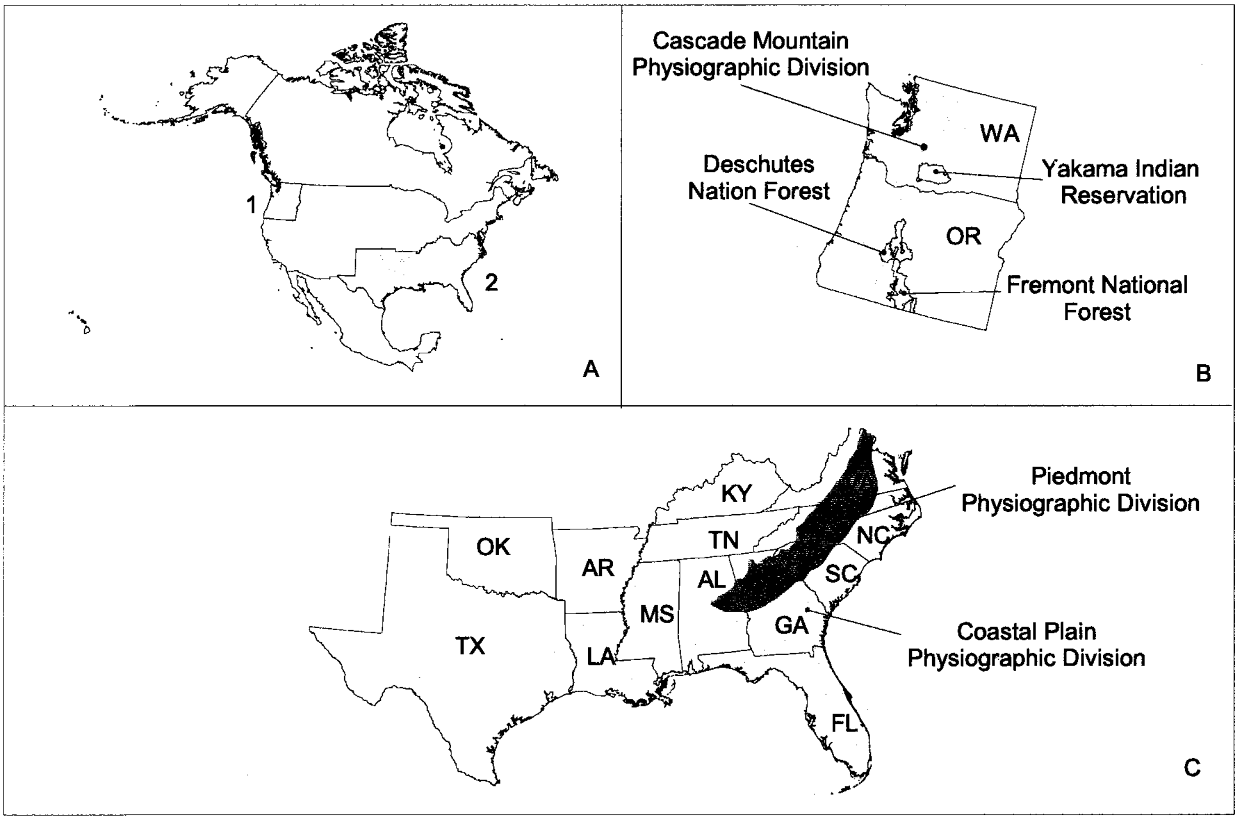

Figure 2. Location of (1) the Pacific Northwest and (2) southeastern United States with respect to North America (A). The Pacific northwestern United States (B) is comprised of Washington (WA) and Oregon (OR). The southeastern United States (C) includes: Alabama (AL), Arkansas (AR), Florida (FL), Georgia (GA), Kentucky (KY), Louisiana (LA), Mississippi (MS), North Carolina (NC), Oklahoma (OK), South Carolina (SC), Tennessee (TN), Texas (TX), and Virginia (VA).

tion is an indicator of both biological diversity and forest health and vitality. Disturbances such as insect and pathogen activity are indicators of forest health and vitality. Spatial scan statistics are potentially applicable to other indicators of both ecological and forest health.

Spatial Cluster Analysis of Forest Fragmentation Patterns

Forest perforations are an important forest health issue because they introduce edge effects near the interior of forest patches, and the resulting edge effects are disproportionate to the area that has been fragmented (Reed and others 1996). While some degree of perforation is normal in all forests (e.g., storm damage), large numbers of perforations associated with urban and agriculture land-cover types are not normal. Some species and ecological processes may benefit from anthropogenic forest perforations, but a forest health issue arises if those species and processes do not represent natural conditions. Knowledge of the location of geographic clusters of perforated forest is useful because perforations tend to grow and coalesce over time, such that clusters of perforated forest may represent large regions close to a transition to a predominantly patchy forested environment (Wickham and others 1999).

This example uses a data set comprised of county-level forest fragmentation statistics that were compiled for forested areas of a 13-state region in the southeastern United States. The statistics were computed from land-cover maps derived from 1992 satellite imagery (Vogelmann and others 2001). Using procedures described by Riitters and others (2002), each of the approximately 993 million 0.09-ha units of forestland in the study area was assigned a label indicating the type of forest fragmentation, if any, observed in the surrounding 7.29-ha neighborhood. Approximately 201 million units were classified as perforated forestland, defined as

Table 1. Summary statistics for forest fragmentation and insect and pathogen data sets

Example	Number of measurement units (i)	Total population (M^T)	Total number of events (N^T)	Annual events/100,000
Forest fragmentation	1072	893944(km^2)	180615(km^2)	20217.7
Insect and pathogen activity	7376	19803396 (ha)	7406684 (ha)	3740.5

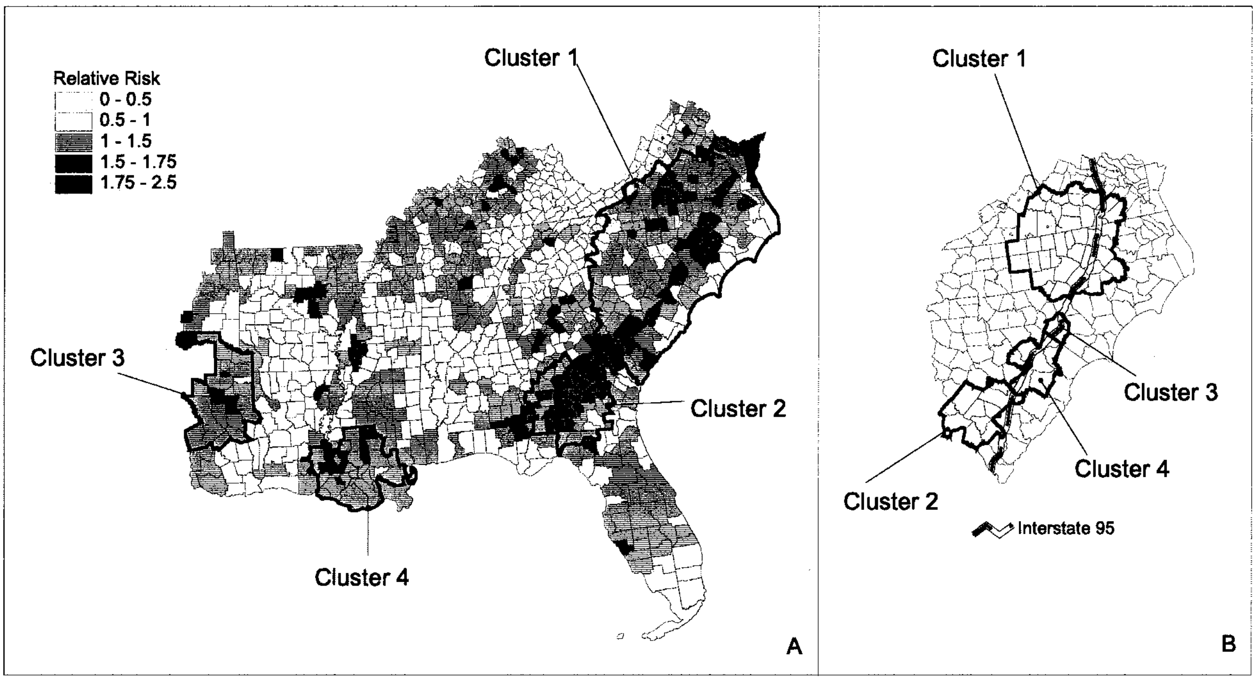

Figure 3. Relative risk, or the ratio of the observed to the expected proportion of forestland in each county that has a perforated forest label and the four most likely regional clusters of perforated forest conditions (A). Local hotspots of perforated forest conditions within the largest regional cluster (B).

a unit of forestland that exists within a 7.29-ha landscape that is more than 60% forested and is within 120 m of a nonforest pixel (Riitters and others 2002). Total forestland area, and perforated forestland area were aggregated to the county level to identify clusters of counties with a high proportion of perforated forest.

The data set contains 1072 county records, and each record has the population of 0.09-ha units of forestland in the county (M_i) and the number of events taken to be the number of those 0.09-ha units that had the perforated forest label (N_i) (Table 1). The data were converted to square kilometers, the size of the circular window was permitted to vary from 0 to 50% of the study area, and 9999 replications were used in the Monte Carlo simulation. Figure 3A illustrates the county-level relative risk values (equation 3) for the study area. The visual impression is that the high-risk counties occur in the Piedmont and Coastal Plain regions where forestland is already known from other information to be juxtaposed with urban and agricultural land uses. It appears that there may be several clusters in that area, but counties with high relative risk also occur elsewhere in the study area and it is not visually apparent whether or not these are clustered.

The four most likely clusters found by the spatial scan statistic (Table 2) are shown in Figure 3A. The two most probable clusters of counties identified by the scan statistic include most of the high-risk counties in the Piedmont and Coastal Plain regions (Figure 3A). Cluster 1 includes all of the high-risk counties in Virginia, North Carolina, and South Carolina, and cluster 2 includes many of the high-risk counties in southern Georgia. Clusters 3 and 4 contain mostly medium- to high-risk counties in eastern Texas and southern Louisiana and Mississippi, respectively.

Table 2. Summary and test statistics for spatial clusters of forest fragmentation and spatiotemporal clusters of insect and pathogen activity

Example	Population (M_c)	Number of events (N_c)	Annual events/100,000	Expected number of events (μ)	Log likelihood ratio (Ψ_w)	P
Forest fragmentation						
Cluster 1[a]	142,200	36,556	25,724.5	28,730.5	1185.5	0.0001
Cluster 2	39,784	10,833	27,247.6	8,038.1	460.5	0.0001
Cluster 3	23,221	6,099	26,282.5	4,691.6	198.3	0.0001
Cluster 4	17,302	4,506	26,060.5	3,495.8	136.5	0.0001
Insect and pathogen activity						
Cluster 1[a]	236,458	400,910	33,913.6	44,218.9	535,932.1	0.0001
Cluster 2	257,091	188,468	14,663.3	48,077.4	118,426.4	0.0001
Cluster 3	40,364	46,759	23,171.4	7,548.3	46,167.1	0.0001

[a]Denotes the primary cluster identified by the spatial scan statistic. All other clusters are considered secondary clusters and are ranked by their log likelihood ratio.

When evaluating clusters in relation to the map of relative risk (Figure 3A), it is apparent that seven high-risk counties near the Georgia–South Carolina border fell between clusters 1 and 2 but were not contained in either. If they had been included, then the two clusters would have merged into one. The most likely explanation is that the software considered those seven counties to be part of a secondary cluster that overlapped the most likely cluster, and as a result the cluster was not reported. Furthermore, the largest cluster contains many medium-risk counties, indicating that the actual cluster shape may not be circular. Those observations suggest that reapplying the algorithm over a smaller area will help to resolve spatial clusters.

When the spatial scan statistic was applied to the subset of counties contained in cluster 1, the results (Figure 3B) suggest that cluster 1 is comprised of several smaller clusters arranged in a linear fashion along Interstate 95. This makes intuitive sense because the pattern of land use and development near urban areas and major transportation corridors often produces fragmented forest conditions. This, in turn, suggests that instead of counties, a better aggregation scheme for evaluating regional patterns of forest fragmentation might be based on proximity to urban areas and transportation corridors.

In summary, the scanning procedure found four regional clusters of perforated forest in the southeastern United States, including one regional cluster that was a collection of subregional hotspots. It was clearly useful to apply the scan statistic at multiple scales with this data set, and subsequent analyses might use measurement units other than counties to better locate significant clusters. With this information it will be easier to focus the subsequent assessment of fragmentation impacts on forest health in places where such impacts are most likely to be found.

Space–Time Cluster Analysis of Insects and Pathogens

Insects and pathogens influence forest succession, productivity, and stability through complex ecosystem interactions (Berryman 1986). They are a natural and essential component of forest ecosystems (Castello and others 1995) and are influenced by climate, land management, plant defenses, and predators. Interactions between forests, insects, and pathogens occur at small scales within forest tracts and at large scales over large forested regions. Temporal trends are also of interest because of concern for the cumulative impacts of insects and pathogens over large areas and the possible management and ecological implications associated with such trends. Space–time cluster analysis of insect and pathogen disturbance is useful in several ways. Knowledge of the location of significant clusters can be used to allocate funding at the strategic planning level and to help identify emerging forest pest complexes that may warrant further investigation.

The US Forest Service conducts annual low-altitude aerial surveys to identify damage to forested areas. The damage from insects and pathogens is recorded by making maps while surveying an area from an airplane. The information is used to evaluate overall health conditions and to identify specific areas for possible treatment to reduce the impacts of major outbreaks. Ten years of maps (1990–1999) for the states of Washington and Oregon were used in this example.

The amounts of forestland and forestland with insects or pathogens were first summarized within a lat-

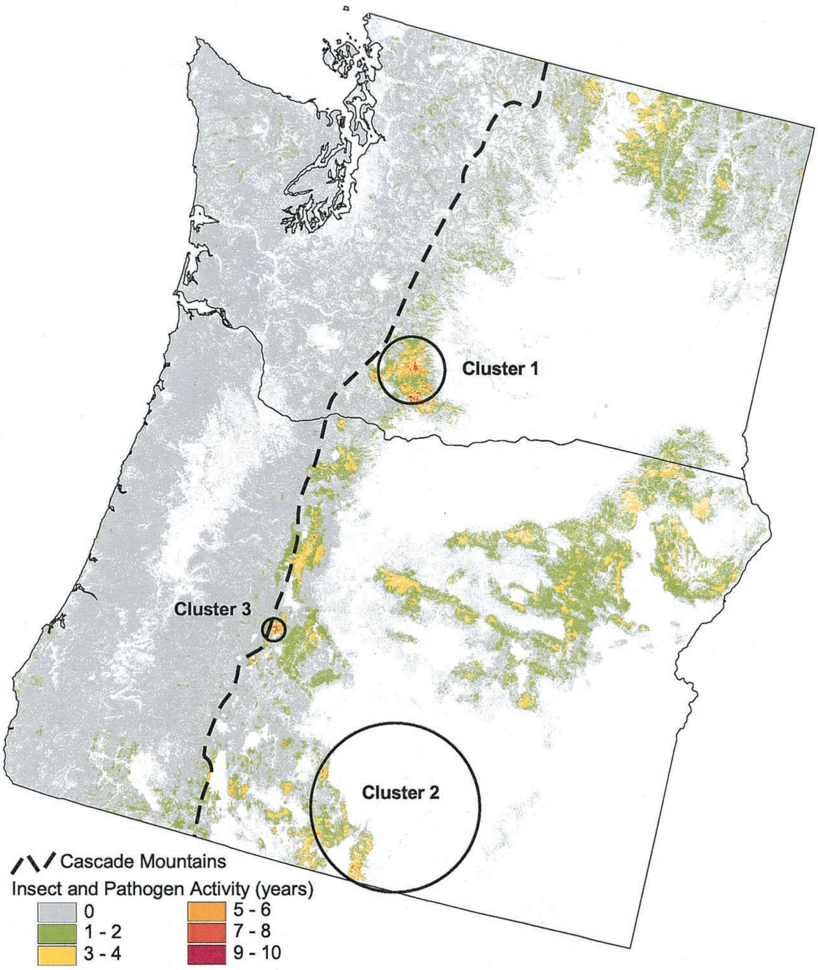

Figure 4. The three most likely current space–time clusters of insect and pathogen presence in Washington and Oregon.

tice of 56.25 km^2 square tiles for Washington and Oregon. The insect and pathogen information was summarized separately for each year of the study and there was no differentiation between specific insects or pathogens. The amount of forestland was based on land-cover maps derived from 1992 satellite imagery (Vogelmann and others 2001), and this measurement was used for all years of the study. Summary statistics are provided in Table 1. The number of hectares of forestland represents the population within each tile (M_i) and the number of hectares with insects or pathogens present represents the events (N_i). As usual, the N_i values are assumed to be Poisson distributed, and tiles are assumed to have equal risk of insect or pathogen activity under the null hypothesis. Under the alternative hypothesis, there exists at least one space–time cluster where the risk inside the cluster was greater than outside. Because forestland is not continuous across the entire study region, the maximum cluster size was fixed at 20% of the total population to ensure that clusters did not contain large unconnected forest areas. The clusters were also forced to contain at least one event from the measurement year 1999 because interest centers on contemporary clusters. For the Monte Carlo simulation, 999 replications were used.

The three most likely space–time clusters identified by the spatial scan statistic are shown in Figure 4 and summary statistics are presented in Table 2. These clusters occurred east of the Cascade Mountains and the most likely cluster was identified in the Yakama Indian Reservation that exhibited a trend of increasing insect and pathogen events during the study period (Figure 5). Within this cluster, there were approximately 179,000 ha of forestland with insects or pathogens present in 1998 (Figure 5). About 94% of the events in the cluster across years were attributable to western spruce budworm (*Choristoneura occidentalis*).

Cluster 2 was in southern Oregon and it included part of the Fremont National Forest (Figure 4). Fir engraver (*Scolytus ventralis*), Modoc budworm (*Choristo-*

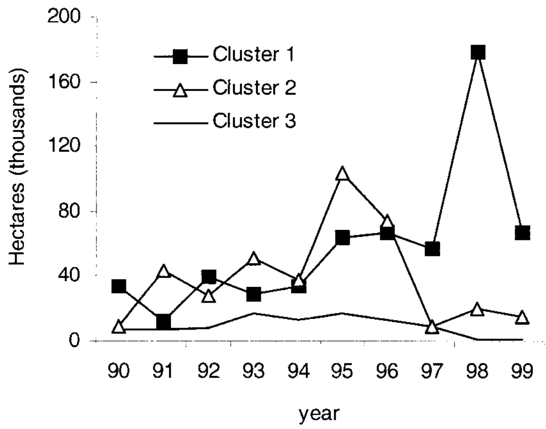

Figure 5. Ten-year trends of forestland area experiencing insect and pathogen presence within the three most likely clusters identified by spatial scan statistics. The values have not been adjusted for the different cluster sizes.

neura retiniana), and mountain pine beetle (*Dendroctonus ponderosae*) accounted for 75.6, 11.3, and 6.5% of the events in the cluster, respectively. The number of events in this cluster increased until 1995 and peaked at approximately 103,000 ha of forestland (Figure 5). Mountain pine beetle was associated with 99% of the events in cluster 3, located in the Deschutes National Forest in central Oregon. The number of events in this cluster also increased from 1990 to 1995 (Figure 5).

The area identified by cluster 1 is part of a well-known western spruce budworm infestation that began in 1984 and is unprecedented in duration in this region. For the past several years, forests in this area were treated with a microbial pesticide (*Bacillus thuringensis*) and thinned in an attempt to curb the infestation. Because the forests in this area have been defoliated for several years, they are likely to be susceptible to secondary insects or pathogens. For example, the douglas-fir beetle (*Dendroctonus pseudotsugae*) is of particular interest because it causes mortality and has already been observed in the area. Considering the region containing the cluster, some possible secondary impacts of the budworm and subsequent infestations might include increased fuel buildup and loss of nitrogen and phosphorous to aquatic systems.

The area identified by cluster 2 is an interesting result and illustrates the usefulness of spatial scan statistics for identifying areas that warrant further investigation. Three different agents accounted for approximately 93% of the activity in the area, however, four additional agents were also present. The shape and location of this cluster were also noteworthy. The cluster was located along the forest grassland ecotone in southern Oregon forcing the shape to be relatively linear (Figure 4). There is no obvious explanation for the location, shape, and agents present in the second most likely cluster and a more detailed analysis may help to explain this result.

Discussion

A comparison of the scan statistic procedure with classical methods further demonstrates the utility of the approach. A classical approach for the forest fragmentation data starts by examining statistics among States. When that is done, the best evidence for a spatial cluster is in the State of Georgia, where the average county has 25.4% of its forest classified as perforated; this is the largest state-level average observed. Lacking any other decision rule, the counties in Georgia with the highest values could be identified. This group comprises counties that roughly correspond with the second most likely cluster from the scan statistic (see Figure 3). The classical method did not identify the largest and most likely cluster because low values in western North Carolina counties obscured the high values in eastern counties.

The classical approach may appear contrived because states should not be used for stratification since there is no reason to expect forest fragmentation patterns to follow state boundaries. However, the same argument applies when aggregating any indicator (e.g., forest fragmentation, air pollution, tree size measurements, etc.). Biophysically defined units such as ecoregions presumably make sense for some indicators, but there is no reason to expect good results for all indicators. Forest fragmentation indicators, for example, exhibit local patterns that are summarized better by county boundaries than by ecoregion boundaries, only because the counties are smaller than ecoregions. The scan statistic makes it possible to identify spatial clusters as groups of neighboring measurement units that may or may not all appear in the same assessment region. Later in the assessment process, individual assessment units could be evaluated based on whether they contain part of a significant cluster.

Depending on the application, it may make sense to look for clusters where the risk of events is less (not more) than elsewhere. For example, from a land management perspective, clusters with low forest perforation values might be candidates for preservation whereas clusters with high perforation would be candidates for mitigation. The examples presented here utilized one-sided tests of the hypothesis and, as a result, only the clusters with significantly more than the expected number of events were detected. The procedure can detect clusters with significantly less than the expected number of events by changing the direction of the inequality in the indicator function in equations 1 and 2. An analysis sensitive to both cluster types requires a two-sided test, which can be performed by dispensing with the indicator function.

Clearly, the analyst must make choices of scale when implementing the scan statistic procedure. There are two general strategies for varying the scales over which clusters are tested. First, the scanning procedure can be tuned to detect clusters of different sizes by limiting the maximum circle size. The forest fragmentation analysis used a 50% (of the total population) upper limit, which meant that the search was optimized for detecting clusters in the interval (0, 50) percent of the total population. The insect and pathogen analysis used a 20% upper limit to help prevent clusters from containing large non-forest areas that are between the mountain ranges in the study area.

If the data and computer resources permit, the second strategy involves a redefinition of the measurement unit. For example, the fragmentation analysis started with data that were already aggregated within county-size measurement units, but the scanning procedure could have instead been directed at much smaller units of aggregation. In that case, subcounty resolution of clusters would have been possible. More recent versions of the SatScan software offer additional options that were not available at the time our study was conducted.

Conclusion

Several national monitoring programs assemble large, spatially explicit databases from field sampling, aerial survey, and satellite imagery. With recent advances in computer hardware and information management, it is now possible to examine forest health issues that were intractable only a decade ago. Creative geographic analyses of time-series data are needed to address a common goal of identifying specific places where the available indicators signal relatively poor forest health conditions. Classical approaches that start by aggregating data according to pre-defined strata may mask forest health problems, or fail to detect problems that span parts of several strata. Based on our experiences with real-world data, spatial scan statistics appear to be a powerful tool for overcoming some of these limitations.

Acknowledgments

The research described in this article was supported in part by a Cooperative Agreement between North Carolina State University and the US Forest Service. Mention of trade names does not constitute endorsement or recommendation for use. We would like to thank Barbara Conkling, Martin Kulldorff, and two anonymous reviewers for their helpful comments and suggestions.

Literature Cited

Berryman, A. A. 1986. Forest insects: principles and practice of population management. Plenum Press, New York 279 pp..

Castello, J. D., D. J. Leopold, and P. J. Smallidge. 1995. Pathogens, patterns, and processes in forest ecosystems. *BioScience* 45:16–24.

Conkling, B. L., Coulston, J. W., and Ambrose, M. A. (eds.). 2003. Forest health monitoring national technical report 1991–1999. US Department of Agriculture, Forest Service, Southern Research Station, Asheville, North Carolina in press.

Czaplewski, R., R. Reich, and W. Bechtold. 1994. Spatial autocorrelation in growth of undisturbed natural pine stands across Georgia. *Forest Science* 40:314–328.

Flather, C., M. Knowles, and I. Kendall. 1998. Threatened and endangered species geography. *BioScience* 48:365–375.

Fotheringham, A. S., and D. W. S. Wong. 1991. The modifiable area unit problem in multivariate statistical analysis. *Environment and Planning A* 23:1025–1044.

Glaz, J., and N. Balakrishnan. 1999. Scan statistics and applications. Birkhauser, Boston 324.

Hjalmars, U., M. Kulldorff, G. Gustafsson, and N. Nagarwalla. 1996. Childhood leukemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine* 15:707–715.

Kulldorff, M. 1997. A spatial scan statistic. Communications in Statistics:. *Theory and Methods* 26:1481–1496.

Kullforff, M., E. Feuer, B. Miller, and L. Freedman. 1997. Breast cancer in northeastern United States: A geographical analysis. *American Journal of Epidemiology* 146:161–170.

Kulldorff, M., W. Athas, E. Feuer, B. Miller, and C. Key. 1998a. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos. *American Journal of Public Health* 88:1377–1380.

Kulldorff, M., Rand, K. Gherman, G., Williams, G. and Defrancesco, D. 1998b. SaTScan v2.1: software for the spatial and space–time scan statistics. National Cancer Institute, Bethesda, Maryland.

Milne, B. T. 1992. Spatial aggregation and neutral models in fractal landscapes. *American Naturalist* 139:32–57.

Naus, J. 1965. The distribution of the size of the maximum cluster of points on the line. *Journal of the American Statistical Association* 60:532–538.

Omernik, J., and R. Bailey. 1997. Distinguishing between watersheds and ecoregions. *Journal of the American Water Resources Association* 33:935–949.

O'Neill, R. V., R. H. Gardner, M. G. Turner, and W. H.

Romme. 1992. Epidemiology theory and disturbance spread on landscapes. *Landscape Ecology* 7:19–26.

Plotnick, R., R. Gardner, and R. O'Neill. 1993. Lacunarity indexes as measures of landscape texture. *Landscape Ecology* 8:201–211.

Reed, R. A., J. Johnson-Barnard, and W. A. Baker. 1996. Contribution of roads to forest fragmentation in the Rocky Mountains. *Conservation Biology* 10:1098–1106.

Riitters, K. H., J. D. Wickham, R. V. O'Neill, K. B. Jones, E. R. Smith, J.W. Coulston, T. G. Wade, and J. H. Smith. 2002. Fragmentation of continental United States forests. *Ecosystems* 5:815–822.

USDA Forest Service. 2000. USDA Forest Service strategic plan (2000 revision). FS-682. US Department of Agriculture, Forest Service, Washington, DC, 73 pp.

USDA Forest Service. 2001. 2000 RPA assessment of forest and range lands. FS-687. US Department of Agriculture, Forest Service, Washington, DC, 78 pp.

USDA Forest Service. 2003. Report of the United States on the criteria and indicators for the sustainable management of temperate and boreal forests. US Department of Agriculture, Forest Service, Washington, DC.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. Van Driel. 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering & Remote Sensing* 67:650–662.

Wickham, J. D., K. B. Jones, K. H. Riitters, T. G. Wade, and R. V. O'Neill. 1999. Transitions in forest fragmentation: implications for restoration opportunities at regional scales. *Landscape Ecology* 14:137–145.