Geographic surveillance of plot-level forest health indicators

John W. Coulston, KaDonna C. Randolph, Anita K. Rose, and William A. Bechtold **US Forest Service, Southern Research Station**

Introduction

Surveillance for the detection of spatial-temporal clusters of elevated risk is common in human epidemiology where the goal is to locate and evaluate potential disease hot spots. The same techniques are applicable to detection monitoring to locate and evaluate potential forest health threats that may occur as hot-spots in space and time. In either human epidemiology or forest health detection monitoring, if a hot-spot is detected and verified during follow-up evaluation, then management decisions can be made to reduce or eliminate the threat.

The objectives of this research were to:

- Examine the Type I and Type II error rates of the spatial scan statistic when using plot-level indicators of forest health risk.
- Describe the landscape characteristics when the the spatial scan statistic using plot-level information has a low Type II error rate.

Background

Characteristics of the spatial scan statistic:

- The procedure is designed to detect clusters and test their significance.
- Cluster sizes and regions do not have to be specified in advance.
- Null and alternative hypotheses are clearly defined.
- The test statistic based on a likelihood ratio, not an *ad hoc* procedure.
- The procedure is valid regardless of the actual spatial pattern.
- The procedure works with data at multiple spatial scales.
- Covariates can be used to account for confounding factors in the background population.
- Scan statistic software is available for free at www.satscan.org

The scanning process procedure:

Procedure starts at measurement unit 1 where a set of concentric circles are generated and the test statistic is calculated for each circle until the maximum circle size is achieved (fig. 1). This process is carried out for each measurement unit. The most likely cluster is the one with the largest test statistic

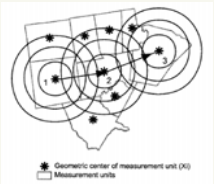

Figure 1. The scanning process.

Using the spatial scan statistic with plots: an example

Suppose we calculate the average crown dieback measured on plots in the South and we consider the 10% of plots with the highest crown dieback to have a forest health risk. Each plot can then be categorized as at risk or not at risk. The spatial scan statistic can then be used on the plots to test for spatial clustering of plots with a forest health risk.

Methods

We used the following simulation approach to determine Type I and Type II error rates:

- Simulate 231 population maps where each 6000 acre pixel was either non-forest, forest, or forest at risk of a health issue. Simulated maps were designed to produce various combinations of percent forest at risk (5%, 10%, 15%), contagion of risk (0 to 1), and range of spatial autocorrelation (20km to 380 km) (fig 2).
- Overlay the standard P3 plot grid onto the simulated maps so that each plot is classified as either non-forest, forest, or forest at risk of a health issue based on the underlying map.
- Run the spatial scan statistic on the 231 population maps at the pixel-level
- Run the spatial scan statistic based on the plot locations for each of the 231 maps

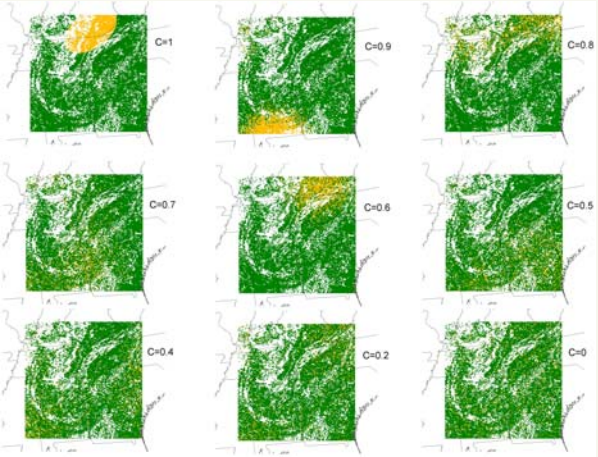

Figure 2. Example of simulated maps. Each of the maps shown here has 10% of the Forest area at risk (orange). The range of spatial autocorrelation is 380km and the Contagion (C) goes from 1 to zero.

The test statistic and null hypothesis:

Within each scanning window the test statistic was calculated and statistical significance was assigned based on Monte Carlo simulation. The null hypothesis (H_0) is that no clustering occurs.

We compared the scan statistic results from the simulated maps (i.e., ground truth) to the scan statistic results based on the plot locations:

- There were 231 comparisons.
- Type I error: Rejecting a true H_0 based on the plots when H_0 was false based on the map (i.e., detecting a cluster that does not actually exist).
- Type II error: Not rejecting H_0 based on the plots when H_0 was true based on the map (i.e., failure to detect an actual cluster).
- Perform analysis of variance to examine the influence of landscape pattern on Type I and Type II error. Landscape variables were:
  - percent and connectivity of risk
  - size of the largest patch of risk and density of risk patches
  - mean and standard deviation of distance between risk patches

Results

Examples of results from individual simulations

The maps in Figure 3 illustrate the following results:

Map (1) Agreement: plot analysis detected significant clusters on the map.

Map (2) Type II error: a significant cluster on the map not detected on plots.

Map (3) Agreement: no significant cluster on the map and none detected on the plots.

Map (4) Results were considered to agree when there was at least 50% overlap between significant clusters identified on both maps and plots.

Map (5) Type I error: a significant cluster was identified on the plots but none existed on the map.

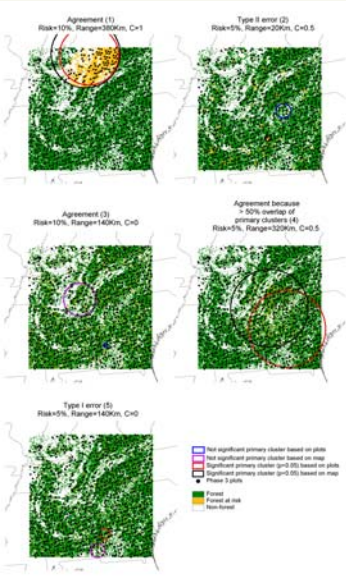

Figure 3. Examples or results from several different simulated maps. Risk is the percent of forest area at risk. Range is the range of spatial autocorrelation in the map. C is the contagion in the map.

Overall Results:

- Type I error rate was 0.4%
- Type II error rate was 15.5%
- The statistically significant variables between the plots and the maps were connectivity of risk, size of the largest patch of risk, and density or risk patches.
- When the largest patch size was greater than 700,000 acres, the type II error rate was approximately 5.5 %.

Conclusions

- Type I errors occur infrequently
- The Type II error rate was relatively high when all patch sizes were included.
- We suggest that the spatial scan statistic can identify spatial clusters of significant risk on the Phase 3 plot grid when the clusters are greater than 700,000 ac.
- Based on the intensity of the Phase 3 sample, we suggest that clusters should contain at least six Phase 3 plots to be considered for follow-up investigation.
- The overall success of correctly identifying areas at risk will also depend on the sensitivity and accuracy of the Phase 3 variable under consideration.

CITATION

Coulston, J.W; Randolph, K.C; Rose, A.K; Bechtold, W.A. 2008. Geographic surveillance of plot-level forest health indicators [Poster Presentation]. Forest Health Monitoring Workgroup Meeting, San Antonio, TX, Feb. 11th – 14th, 2008.

ONLINE AT: http://fhm.fs.fed.us

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Forest Health Monitoring Program for supporting this research.