Data, Data Everywhere: Detecting Spatial Patterns in Fine-Scale Health Data Collected across a Continent

Kevin M. Potter[1], Frank H. Koch[2], Christopher M. Oswalt[3], and Basil V. Iannone[4]

[1] Department of Forestry and Environmental Resources, North Carolina State University, Research Triangle Park, NC 27709 USA; kpotter@ncsu.edu;
[2] Eastern Forest Environmental Threat Assessment Center, U.S. Forest Service, Research Triangle Park, NC 27709 USA; [3] Forest Inventory and Analysis Program, U.S. Forest Service, Knoxville, TN 37919 USA; [4] School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611 USA

Introduction

Fine-scale forest health data are increasingly available across broad regions. Geographic analyses of these data can help identify locations of forest health concern, but it is a challenge to present this information in ways that are relevant for policy and management decisions.

We describe here an approach to identify locations where forest threats occur at greater or lower frequencies than expected by chance (Potter *et al.* 2016).

This method is a standard component of annual national reports on forest health status and trends across the United States.

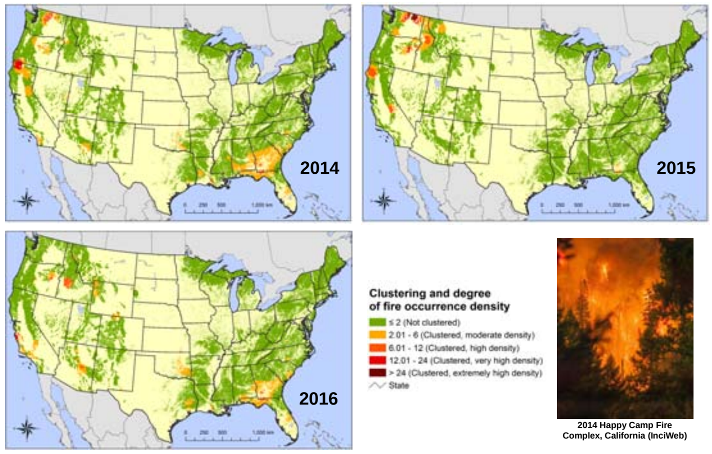

***Figure 2:* Hotspots of forest wildfire occurrence across the United States, from the Moderate Resolution Imaging Spectroradiometer (MODIS) Active Fire Detections database. Values are Getis-Ord G_i* scores for 2014, 2015, and 2016 fire occurrence densities (Potter and Conkling 2016, 2017, in press). Values greater than 2 represent areas of significant clustering of high fire occurrences.**

Results

- The SASH approach identifies statistically significant clusters of forest health threats using broad-scale data sources.
- Hotspots of high fire occurrence densities were detected in northern California and the interior Northwest in 2014, 2015 and 2016 (Figure 2). Lower intensity hotspots were also detected in the Southeast.
- Hotspots of high tree mortality in 2015 were caused by bark beetle species in the West, and by sudden oak death in California (Figure 3). Hotspots in the East were caused by emerald ash borer and gray brown sap rot.

Method Overview

The Spatial Association of Scalable Hexagons (SASH) approach is based on a hexagonal sampling frame optimized for spatial neighborhood analysis, adjustable to the appropriate spatial resolution, and applicable to multiple data types. The method:

- Divides the United States into equal-area hexagonal cells of the appropriate area (Figure 1A) and
- Uses a Getis-Ord (G_i*) hotspot analysis (Getis and Ord 1992) to identify geographic clusters of forested hexagons with high values for the forest threat, such as wildfire (Figure 1B) or insect and disease damage (Figure 1C) in ArcMap® 10.1.

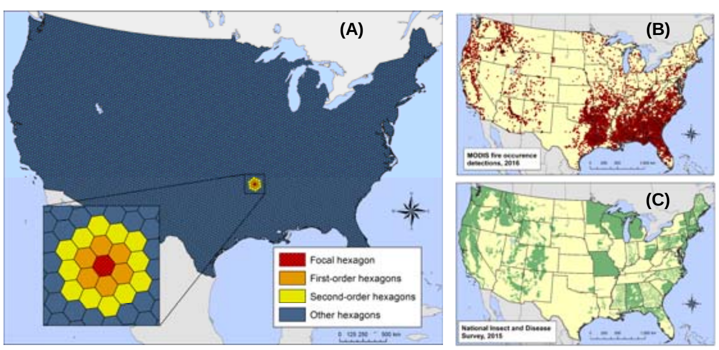

***Figure 1:* (A) hexagonal sampling frame of the continental United States, consisting of 9,810 hexagonal cells of 834 km². The G_i* statistic for each hexagon sums the differences between the mean values in a local sample (inset) and the global mean of all the hexagons in the analysis; data analyzed here are (B) MODIS satellite fire occurrence detections and (C) mortality data from national aerial survey efforts.**

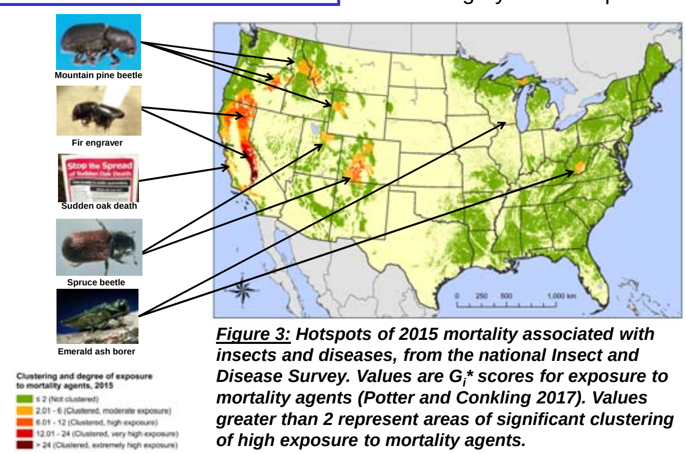

***Figure 3:* Hotspots of 2015 mortality associated with insects and diseases, from the national Insect and Disease Survey. Values are G_i* scores for exposure to mortality agents (Potter and Conkling 2017). Values greater than 2 represent areas of significant clustering of high exposure to mortality agents.**

Discussion

The SASH method is a "big data" analysis tool useful for ecological studies that require rigorous testing of hypotheses within a spatial framework. It can be applied easily across many regions and datasets.

It is useful for understanding macroscale patterns and processes associated with forest health threats, and for identifying high-impact areas where specific management activities may be needed.

Method Details

Forest health data are aggregated to hexagons, which are compact and uniform in their distance to the centroids of neighboring hexagons (useful attributes for a spatial neighborhood analyses).

The hotspot statistic G_i* (Figure 1A) is a standardized *z*-score with a mean of 0 and a standard deviation of 1, with values > 1.96 indicating significant clustering of high threat values ($p < 0.025$) and values < -1.96 indicating significant clustering of low values.

Data analyzed here are **(1)** forest fire occurrences detected daily by MODIS satellite for 2014, 2015, and 2016 (Figure 1B) and **(2)** forest mortality caused by insects and diseases, from 2015 national aerial survey data (Figure 1C). For the analysis, all are standardized by the amount of forest area within each hexagon.

Application

SASH analyses (of forest fire occurrences and mortality/ defoliation caused by insects and disease) are presented annually in ***Forest Health Monitoring: Status, Trends and Analysis*** reports published by the U.S, Forest Service, available online.

https://fhm.fs.fed.us/pubs

References

Getis, A., J.K. Ord. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis*. 24(3):189-206.

Potter, K.M., F.H. Koch, C.M. Oswalt, and B.V. Iannone. 2016. Data, data everywhere: Detecting spatial patterns in fine-scale ecological information collected across a continent. *Landscape Ecology.*31:67-84. DOI: 10.1007/s10980-015-0295-0.

Potter, K.M., and B.L. Conkling, editors. 2017. Forest Health Monitoring: National Status, Trends and Analysis, 2016. General Technical Report SRS-222. Asheville, North Carolina: U.S. Department of Agriculture, Forest Service, Southern Research Station. 195 p.

Potter, K.M., and B.L. Conkling, editors. 2016. Forest Health Monitoring: National Status, Trends and Analysis, 2015. General Technical Report SRS-213. Asheville, North Carolina: U.S. Department of Agriculture, Forest Service, Southern Research Station. 199 p.

This research was supported in part through Cost Share Agreements 14-CS-11330110-042, 15-CS-11330110-067, and 16-CS-11330110-038 between the United States Department of Agriculture (USDA) Forest Service and North Carolina State University.