Spatio-Temporal Modeling of Southern Pine Beetle Outbreaks with a Block Bootstrapping Approach

Xuan Chen

Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC 27695, e-mail: xchen8@ncsu.edu

Barry K. Goodwin
William Neal Reynolds Distinguished Professor,
Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC 27695

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011

Spatio-Temporal Modeling of Southern Pine Beetle Outbreaks with a Block Bootstrapping Approach

Xuan Chen and Barry Goodwin

Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC 27695

Abstract

Our study focuses on modeling southern pine beetle (SPB) outbreaks in the southern area. The approach is to evaluate SPB outbreak frequency in a spatio-temporal framework. A block bootstrapping method with zero-inflated estimation has been proposed to construct a statistical model accounting for explanatory variables while adjusting for spatial and temporal autocorrelation. Although the bootstrap (Efron 1979) method can handle independent observations well, the strong autocorrelation of SPB outbreaks brings about a major challenge. Motivated by bootstrapping overlapping blocks method in autoregressive time series scenario (Kunsch 1989) and block bootstrapping method of dependent data from a spatial map (Hall 1985), we have developed a method to bootstrap overlapping spatio-temporal blocks. By selecting an appropriate block size, the spatial-temporal correlation can be eliminated. The second challenge arises from the fact that the SPB spots distribution has a heavy weight on 0. To accommodate this issue, the zero-inflated models are adopted in the estimation stage. With our saptio-temporal block bootstrapping approach, impacts of environmental factors on SPB outbreaks and implications of pine forest management are assessed. Almost all the explanatory variables, including drought, temperature, forest ecosystem and hurricane, have been detected to have significant impacts. Forestland size and government share of forestland would positively contribute to SPB outbreaks significantly. Meanwhile, our method offers a way to forecast the frequency of future SPB outbreaks, given the current environmental information of a county.

Introduction

The southern pine beetle (Dendroctonus frontalis Zimmermann) is one of most destructive insects in the U.S. forest sector (Gan 2001). This topic is of growing interest especially when southern pine beetle (SPB) outbreaks have increased from undetectability in 1960s to epidemics in recent years (Mawby and Gold 1984; Gumpertz 2000). The economic loss during a SPB epidemic season can be as high as millions of dollars (Pye et al.).
Research on SPB usually falls into two groups. The first category focused on the biological behavior of SPB in an ecosystem, and the related literature includes McNulty (1998). The second category focused on SPB outbreaks in across regions, e.g. counties (Gumpertz 2000; Zhu 2005) or states (Gan 2001).

Although SPB risks at county level in the southern states have already been studied with spatio-temporal models (Gumpertz 2000; Zhu 2005), the frequency of SPB outbreaks has not yet been investigated. Most existing literature focused on whether a SPB outbreak would occur in a county no matter how big or small the potential damage is. In other words, a modest endemic and a destructive epidemic were treated indifferently. However, the economic costs vary dramatically when the SPB activity fluctuates (Pye et al.). Hence it is necessary to assess the annual SPB outbreak counts in each county, and the approach of this chapter is to evaluate SPB outbreak frequency in a spatio-temporal framework.

In our study, a block bootstrapping method with zero-inflated estimation has been proposed to construct a statistical model accounting for explanatory variables while adjusting for spatial and temporal autocorrelation. Although the bootstrap (Efron 1979) method can handle independent observations well, the strong autocorrelation of SPB outbreaks brings about a major challenge. Motivated by bootstrapping overlapping blocks method in autoregressive time series scenario (Kunsch 1989) and block bootstrapping method of dependent data from a spatial map (Hall 1985), we have developed a method to bootstrap overlapping spatio-temporal blocks. By selecting an appropriate block size, the spatial-temporal correlation can be eliminated. The second challenge arises from the fact that the SPB spots distribution has a heavy weight on 0. To accommodate this issue, the zero-inflated models are adopted in the estimation stage.

Methods

• ***Bootstrapping (Efron 1979)***

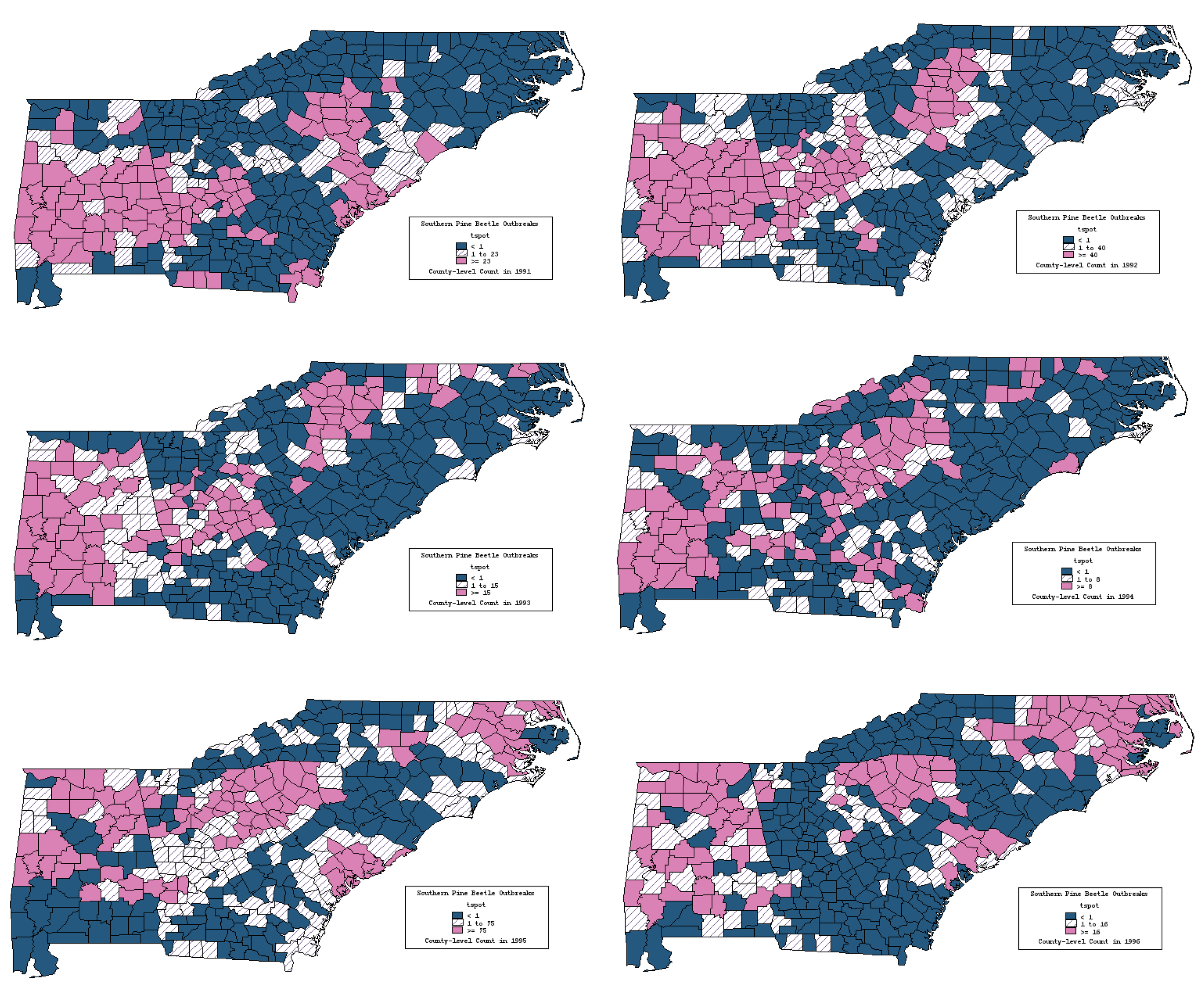
Suppose that $\hat{\theta}_n$ *is* an estimate of a parameter vector θ based on a sample $X= (x_1, x_2 \ldots x_n)$. An approximation to the statistical properties of $\hat{\theta}_n$ can be obtained by studying a sample of bootstrap estimators $\hat{\theta}(b)_m$, $b=1, 2, \ldots .B$, obtained by sampling n observations, with replacement, from X and re-computing $\hat{\theta}$ with each sample. After a total of B times, the desired sampling characteristic is computed from $\Theta = [\hat{\theta}(1)_m \ldots \hat{\theta}(B)_m]$

For example, if it were known that the estimator were consistent and if n were reasonably large, then one might approximate the asymptotic covariance matrix of the estimator

$$\text{Est.Asy.Var}[\hat{\theta}] = \frac{1}{B}\sum_{b=1}^{B}(\hat{\theta}(b)_m - \hat{\theta}_n)(\hat{\theta}(b)_m - \hat{\theta}_n)'$$

Basic Assumptions: Independence

• ***Spatio-Temporal Correlated SPB outbreaks (NC, SC, GA & AL Selected Years)***



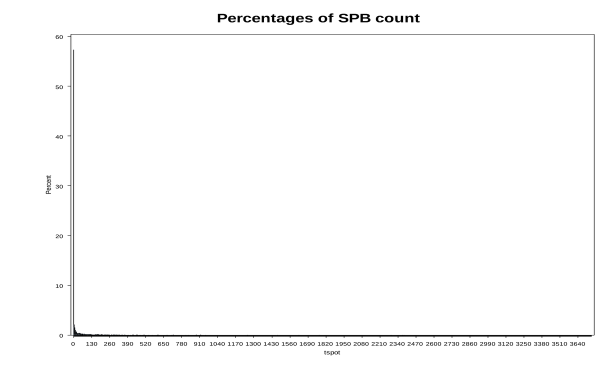
• ***Block Bootstrapping***

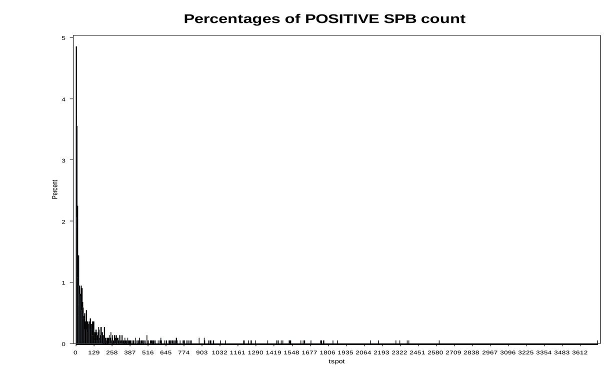
Hall (1995) discussed the technique to deal with dependent data on a spatial map in the bootstrapping context. Motivated by this method, we derived a spatio-temporal block bootstrapping method in the hope to eliminate the spatio-temporal autocorrelation.

First of all, we will divided all the county-year into several blocks. The grouping rule is as follows: if a county at time t , denoted as i, is not at the edge of this spatial-temporal space, then it can be a center of a block. Then all the counties whose centroids are less than 50 miles away from county i's centroid should be thought as spatial neighbors. Moreover, the observations of same county in two adjacent time periods are deemed as temporal neighbors. A block centered at the observation of county i at time t contains all the records of county 's spatial neighbors and their temporal neighbors. In this way, we divided the space-time records into n overlapping blocks. Each iteration, we resample the blocks, with replacement n times.

Estimation

• ***Outbreaks Distribution***





• ***Zero-inflated Exponential Model***

In the first regression stage, logit selection mechanism is used to discriminate the 0s and positive find.

In the second stage, the positive find is fitted to an exponential distribution.

• ***Estimation Results***

Independent Variable	Logit Selection Mechanism for Positive Spots (zero-inflated portion)	Exponential Model of Positive Spots
Intercept	1.43781*** (0.09863)	3.37195*** (0.11654)
Forestland Size (in 1000 acres)	-0.46947*** (0.01298)	0.44226*** (0.01741)
Government owned share of forestland	-0.49071*** (0.16879)	0.95105*** (0.14039)
Longleaf / slash pine & Oak / gum / cypress share	4.20431*** (0.14654)	-2.22207*** (0.17288)
loblolly / shortleaf pine & Oak / pine share	-0.78250*** (0.11238)	0.62992*** (0.11449)
Two years' daily average of CDD Index in	-0.14536*** (0.02157)	-0.04243 (0.02652)
Two years' monthly average of PHDI Index	0.13341*** (0.01268)	-0.22237*** (0.01682)
Hurricane incidences	0.83123*** (0.12094)	-0.65892*** (0.13438)

Note: Iteration=100. Single (*), double (**), and triple (***) denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Conclusion

The scale factor, forestland size and government share of forestland would positively contribute to SPB outbreaks significantly. The reason that SPB spots are more likely to happen on federal/state owned forest field is not yet clear. One possible reason is that private owners would pay more attention toward their properties. The other possible explanation stems from the fact the public forestlands are large in size which makes the SPB contamination easy to happen.

Drought is found to enhance SPB outbreaks while hot weather has a mixed impact. First of all, we used the average drought index and CDD index (measurement of hot weather) in two years because the life cycle of SPB is two years (Ungerer et al. 1999). PHDI index, with a lower value equivalent to drier ground condition, has a negative sign in the positive count. It means the drought is helpful for SPB to survive. The hot weather could significantly reduce the chance that SPB outbreaks will not happen. However, its impact on the positive find is in a reverse way, though not significant.

Hurricanes, usually predicted to help with SPB epidemics, could reduce SPB outbreaks in our results. A potential explanation would be that lying timbers knocked down by the hurricanes are removed quickly after disasters. Hence, instead of being provided bridges to move around, SPB schools in different locations would be separated.

Among the forestland ecosystem, loblolly / shortleaf pine & oak / pine forestland would enhance SPB outbreaks while longleaf / slash & oak / gum / cypress forestland would reduce SPB outbreaks. As the host type for SPB, a higher proportion of loblolly / shortleaf pine & oak / pine forestland will help with more SPB outbreaks.