Characterizing Spatial Patterns of Gypsy Moth Regional Defoliation

ANDREW M. LIEBHOLD
JOSEPH S. ELKINTON

ABSTRACT. Aerial sketch maps of defoliation caused by the gypsy moth, Lymantria dispar (L.), in Massachusetts from 1961 to 1986 were digitized into an array of 1.9 x 1.9 km cells. Defoliation frequency was mapped by summing defoliation incidence for each cell. Spatial autocorrelation analysis indicated that defoliation was not distributed randomly over the state but was instead distributed in clumps of varying size. Directional spatio-temporal autocorrelation analysis showed no evidence for unidirectional spread of defoliation through time. Yearly defoliation, as a proportion of maximum, was used in a clustering algorithm to categorize different portions of the state in terms of defoliation synchrony. Two-state transition models were used to characterize defoliation initiation and termination. Different regions showed different probabilities of both outbreak initiation and termination. Transition probabilities varied substantially relative to the amount of defoliation in adjoining areas, indicating an apparent "spread" of defoliation. FOR. Sci. 35(2):557-568.

ADDITIONAL KEY WORDS. Spatial autocorrelation, Lymantria dispar, spatial dynamics, population dynamics.

Since its introduction in 1869 near Boston, MA, the gypsy moth, Lymantria dispar (L.), has expanded its range as far as Canada, Michigan, and North Carolina. In many of the areas where gypsy moth is established, populations have not stabilized permanently at low densities. Instead, populations often fluctuate through changes of several orders of magnitude in density.

Little is known about spatial patterns of gypsy moth outbreak development. Valentine and Houston (1979) and Campbell (1976) hypothesized that outbreaks may be initiated by immigration of larvae from nearby outbreak areas. Campbell (1973, 1976) and Campbell and Sloan (1978) concluded that outbreaks are likely to persist from one year to the next when insect densities range widely among subpopulations in a region and conversely that when numerical variability is minimal among subpopulations, outbreaks are likely to decline. Mason and McManus (1981) qualitatively related changes in yearly defoliation patterns in New Jersey to expected first instar dispersal.

Many state and federal agencies are developing new approaches for management of gypsy moth impacts. At the core of modern gypsy moth integrated pest management systems are geographically based monitoring systems (Ravlin et al. 1987, Reardon et al. 1987). The approach is to detect new, rising populations and then treat these populations in order to prevent...
regional outbreaks. In order to evaluate treatment decisions, there is a need for models that predict the large-scale spatial dynamics of gypsy moth populations.

The objective of this study was the development of statistical models that describe the large-scale spatial dynamics of gypsy moth defoliation. These models were used to quantify the extent of aggregation of defoliation through space and time, to explore the extent to which defoliation spreads, and to identify geographic regions where defoliation occurs in synchrony. We expect that an understanding of these spatial processes will eventually aid in the management of gypsy moth outbreaks on a large-scale geographic level.

DESCRIPTION OF THE DATA

The Massachusetts Department of Environmental Management monitors gypsy moth defoliation annually in all parts of the state using aerial sketch maps. Maps are sketched during a series of flights over the state in late July when defoliation is at its peak. 30% defoliation is considered the lower threshold for detection. In situations where there is doubt as to the cause of the defoliation, ground checks for the presence of gypsy moth life stages are made. These maps are initially overlaid over standard U.S. Geological Survey (1:24,000) topographical maps and then used to generate a single 1:760,320 composite map which covers the entire state.

Composite maps for the years 1961–1986 were digitized into a 65 x 101 matrix of cells. Each cell represented a 1.9 x 1.9 km area. The matrix covered all portions of the state. Each cell was coded as either 1: containing defoliation, 2: undefoliated, 3: not part of the state. Maps were digitized by placing a plastic overlay over each composite map and transcribing the value of each cell into a computer file. There is unavoidable error generated in the process of recording defoliation on aerial sketch maps (Talerico 1981). By using large cells (1.9 x 1.9 km) for digitizing defoliation, error in defoliation location was minimized.

Defoliation frequencies for each cell were summed from 1961–1986 (Figure 1). From the original 65 x 101 matrix of defoliation frequencies, a 25 x 25 matrix was interpolated as the weighted mean of the eight closest grid points using an inverse-distance weighting function (Sampson 1975):

\[ w = \frac{(1 - D/(1.1 \times D_{\text{max}}))^2}{(D/(1.1 \times D_{\text{max}}))^2}, \]

where

- \( w \) = weight attached to a sample data point a distance \( D \) from the grid intersection being estimated and
- \( D_{\text{max}} \) = distance from the grid intersection to the most distant of the eight closest original points.

This smaller grid matrix was generated for two reasons: it served to smooth the surface and secondly it reduced the amount of computer memory required to generate contour intervals. The new grid matrix was then used in a piecewise Bessel interpolation algorithm to generate defoliation frequency contour intervals (Figure 2). Portions of Cape Cod and an area near the northeastern border of Rhode Island exhibited the greatest numbers of years of defoliation.
SPATIAL AUTOCORRELATION ANALYSIS

Spatial autocorrelation analysis is a statistical technique developed by statistical geographers for describing the spatial variation exhibited by the response surface of a given variable (Sokal and Oden 1978, Reed and Burkhart 1985). A given spatial autocorrelation statistic quantifies the degree of association of a response variable with spatially adjacent points. Spatial adjacency can be defined in a variety of ways (Cliff and Ord 1973, Sokal and Oden 1978). In this study, we used the "Queen's move" adjacency definition (Figure 3). Under this definition, two points are considered adjacent if
their x or y coordinates differ by a given distance (termed a "distance class"). Typically, separate autocorrelation statistics are calculated for sets of points that are spatially adjacent at different distance classes. Several different autocorrelation statistics are available and are appropriate for different types of response variables. In our situation, the defoliation associated with a given cell was a binary variable (1 if it was defoliated, 0 otherwise). For binary data the appropriate autocorrelation statistic is the join-count statistic:

$$1:1 \text{ Joins} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}x_i x_j$$

where

1:1 Joins = the number of adjacent cells that were both defoliated,

$x_i$ and $x_j$ = the values of the response variable (defoliation status) at points $i$ and $j$ respectively,

$w_{ij}$ = the adjacency weight between points $i$ and $j$ (1 if adjacent, 0 otherwise).

Join-counts were expressed as standard normal deviates so as to remove the direct effect of yearly defoliation frequency (Sokal and Oden 1978).

Figure 4A displays the yearly 1:1 join-count statistics for distance classes of 1 to 5 cells (1.9–9.5 km). Neighboring cells were always significantly autocorrelated (i.e., join-count standard normal deviates for the distance class of 1.9 km were always greater than 2; $\alpha = 0.05$) (Figure 4A). Cells that were short distances apart were usually autocorrelated as well. This showed that defoliation was not randomly distributed through space, but more frequently occurred as clumps of two or more cells. The magnitude of spatial autocorrelation generally paralleled trends in total defoliation (Figure 4). In 1975 there was an unusual drop in autocorrelation. This reflected the distri-
bution of defoliation in several scattered localities following a decrease in defoliation from 1974.

Spatio-temporal autocorrelation analysis (STAA) was also applied to the defoliation data. STAA is used to evaluate the degree of association between adjacent cells where adjacency is described in both space and time. Results from STAA can be useful for understanding the spread of epidemics through space (Cliff and Ord 1973, Reynolds et al. 1988). In this study we took the $65 \times 101 \times 26$ matrix of binary defoliation incidence (2 spatial and

**FIGURE 3.** A. Adjacency definition for a Queen’s case and distance class of 1. B. Adjacency definition for a Queen’s case and distance class of 2.

**FIGURE 4.** Temporal patterns of defoliation autocorrelation and frequency from 1961–1986. A. Normalized defoliation join-counts for adjacent cells which are both defoliated (Sokal and Oden 1978). B. Yearly percent of cells defoliated in Massachusetts.
one temporal dimensions) and calculated join-count statistics using adjacency definitions based on distance classes of 1 through 10 simultaneously through time and space (1.9 to 19 km and 1 to 10 years). We restricted adjacency definitions to separate cardinal directions (Oden and Sokal 1986) (Figure 5). This directional STAA was thus used to statistically evaluate the extent to which defoliation spread in one cardinal direction. Qualitative analysis of other gypsy moth defoliation maps suggested that changes in the spatial distribution of defoliation between years is often unidirectional due to the wind-born dispersal of first instars (Anderson and Gould 1974).

Directional STAA yielded nearly identical correlograms in all four cardinal directions (Figure 6). Cells within six time-space cells were significantly autocorrelated (i.e., join-count standard normal deviates were greater than 2; α = 0.05). Beyond nine units there was evidence of negative autocorrelation. These results suggest that defoliation in a given cell may "spread" to spatially adjacent cells in the future. However, they indicate that there was not an overriding directional component to the spread of the defoliation.

**CLUSTER ANALYSIS**

Cursory examination of defoliation data showed that temporal defoliation patterns were not synchronous among all portions of the state (Figure 1). Cluster analysis was used to categorize regions according to synchrony of defoliation.

All map cells were assigned to one of 30 regions, which consisted of 10 × 10 groups of adjoining cells (Figure 7). The aggregation of the original grid matrix into 10 × 10 groups was performed for two reasons: (1) we intended to identify spatial synchrony only on a relatively large scale, and (2) we were able to generate a continuous variable (yearly defoliation magnitude) which was more readily applied in cluster analysis than the original binary yearly defoliation condition. For each region, yearly defoliation magnitude was the proportion of cells defoliated, expressed as a proportion of the maximum observed defoliation for the region from all years. This relative magnitude was used to eliminate the effect of areas that were incapable of supporting defoliating populations (e.g., agricultural land and resistant stands). Ward's cluster algorithm (Hair et al. 1987) was then used to separate regions according to their squared Euclidean distance in the 26 variables (years). In other words, each region was represented by a point in a

![Figure 5](https://academic.oup.com/forestscience/article/35/2/557/4642609)
26-dimensional space, where each dimension represented the defoliation magnitude for 1 year, and regions were assigned to clusters based on their proximity in this space.

The first cluster to be identified (i.e., most distinctive defoliation pattern in terms of outbreak synchrony) from the body of other regions consisted of
Cape Cod and the southeastern portion of the state (Figures 8 and 9). In this cluster, regions characteristically experienced high defoliation from 1980–1986 and had other more minor eruptions from 1963–1978 (Figure 10). The cluster algorithm identified several other clusters of regions having similar defoliation frequencies though they were not as distinct as cluster 1 (Figures 8 and 9). There are no satisfactory methods for determining the number of population clusters for any type of cluster analysis (Everitt 1980). Therefore, we arbitrarily selected the first five clusters for use in other analyses.

TRANSITION MODELS

Transition models are tools for describing temporal transitions of a variable from one nominal value to another. Transition models have been used as powerful methods for characterizing outbreak patterns from historical data (Kemp 1987).

In this study, we used a two-state transition model (Parzen 1962) to quantify the transition of cells (the original 1.9 x 1.9 km cells) from 0 (undefol-
liated) to 1 (defoliated) and from 1 to 0. Separate transitions probabilities were calculated to compare cells in different clusters (previously identified using cluster analysis). We performed this analysis both ignoring and taking into account the state of adjacent cells [immediate neighbors under a Queen’s move definition of adjacency (Cliff and Ord 1973)].

Simple transition frequencies for the transition $0 \rightarrow 1$ were relatively

---

**FIGURE 9.** Map of clusters generated using Ward’s algorithm.

**FIGURE 10.** Yearly defoliation frequencies in each of 30 regions.
lower for clusters 1, 2, and 5 than for 3 and 4 (Table 1). However, the rank order of $1 \rightarrow 0$ transition probabilities was 3, 1, 4, 5, 2. In other words, the higher defoliation frequencies in clusters 3 and 4 (Figures 2 and 9) can be partially explained by their higher defoliation initiation probability. However, even though cluster 1 had a high defoliation frequency (Figures 2 and 9), it did not initiate defoliation more often than the low defoliation clusters, 2 and 5. The high defoliation frequency of cluster 1 must be due to its low probability of defoliation termination (Table 1). The low defoliation frequency of cluster 5 can be explained by both a low $0 \rightarrow 1$ probability and a high $1 \rightarrow 0$ probability. In contrast, the high defoliation frequencies of clusters 3 and 4 can be explained by both high $0 \rightarrow 1$ probabilities and by low $1 \rightarrow 0$ probabilities.

All clusters exhibited similar trends in defoliation transition frequencies with respect to the defoliation status of adjoining cells. Probabilities of defoliation initiation increased as the proportion of defoliated neighbors increased (Figure 11a). Similarly, the probability of defoliation termination decreased as the proportion of defoliated neighbors increased (Figure 11b). The inverse relationship between $1 \rightarrow 0$ probability and proportion of defoliated neighbors does not contradict Campbell’s (1978) theory that outbreaks are often maintained by redistribution of insects from neighboring outbreak areas.

**DISCUSSION**

These analyses indicate that different regions differ in their defoliation timing and frequency. Aside from the 1981 outbreak, there was little synchrony of defoliation across the state. Southeastern Massachusetts was the most distinct in terms of its defoliation timing. This difference may be explained by climatological, forest stand, or other unknown differences among the regions that affect population processes.

Spatial autocorrelation analysis showed that defoliation is not distributed randomly in space. Instead, defoliation exists as clumps of varying size in space. The development of these clumps through time can be seen in both the STAA correlograms (Figure 6) and the transition frequencies (Figure 11). The existence of neighboring defoliated cells increases the probability of a cell becoming defoliated. This represents what is referred to as the "focal area phenomenon," where visible defoliation begins at some localized area and then this area grows larger through time (Chugunin 1949, Val-
Figure 11. Transition probabilities of transition to a different defoliation state with varying numbers of defoliated neighbors. A. Transition from nondefoliated to defoliated; B. Transition from defoliated to nondefoliated.

Entine and Houston 1979, Wallner 1987). The nearly identical STAA correlograms for each of the cardinal directions (Figure 6) indicated that defoliation spread outward equally in all directions rather than spreading primarily in a single direction as would be expected from the windborne dispersal of larvae in a unidirectional wind (Anderson and Gould 1974).

The outbreak that occurred in 1981 was an unprecedented phenomenon in terms of the overall quantity of defoliated cells and in terms of the high synchrony of most regions during this period. It is interesting to note that many cells become defoliated in the absence of defoliated neighbors (Figure 1). Thus, during this most intense outbreak, the focal area phenomenon was not prominent.

LITERATURE CITED


