

Fragmentation profiles for real and simulated landscapes

GLEN D. JOHNSON,¹ WAYNE L. MYERS,²
GANAPATI P. PATIL¹ and CHARLES TAILLIE¹

¹*Center for Statistical Ecology and Environmental Statistics, Department of Statistics*

²*School of Forest Resources and Environmental Resources Research Institute,
Penn State University, University Park, PA 16802 USA*

Received January 1999; Revised March 2000

When a natural landscape is represented by a series of categorical raster maps of varying resolution, a multiresolution characterization of spatial pattern can be obtained in which entropy is computed at each resolution conditional on the next coarser resolution. The series of entropy values is plotted as a function of resolution, resulting in a multiresolution profile of fragmentation pattern in the landscape.

If a categorical raster map is available at a single resolution only, a series of degraded maps at increasingly coarser resolutions is generated and the fragmentation profile is computed for this series. An algorithm has been developed for obtaining the profile directly from the single resolution map without having to generate and store the coarser resolution maps.

A hierarchical stochastic model is described for simulating categorical raster maps and the fragmentation profile of the generating process is obtained in terms of the model parameters. These “process” profiles provide benchmarks for assessing empirical profiles obtained from raster maps of actual landscapes. Methods of the paper are applied to several watersheds of Pennsylvania using landcover maps derived from satellite imagery. These examples indicate that characteristic landscape types induce characteristic features in their fragmentation profiles.

Keywords: categorical raster maps, conditional entropy, diagonal dominance, HMTM model landscape ecology, Markov transition matrix, multiresolution spatial pattern, self-similarity

1352-8505 © 2001  Kluwer Academic Publishers

1. Introduction

An emerging theme in the field of landscape ecology is the multiresolution characterization of spatial pattern in classified maps such as maps of land use or of landcover. Maps in the latter class are often obtained by classifying satellite images, thus defining the grain of measurement as the pixel size. Landscape patterns have been observed to depend on both the grain and geographic extent of a study area (Wiens, 1995); therefore, we may not want to be restricted to a fixed measurement scale. Observations of scale dependence are not surprising, given that hierarchically-nested spatial patterns have been observed in actual landscapes (Kotliar and Wiens, 1990; O’Neill, Gardner and Turner, 1992) and are expected from ecological hierarchy theory (O’Neill, Johnson and

1352-8505 © 2001  Kluwer Academic Publishers

King, 1989). Ultimately, concern should not lie with determining an appropriate measurement scale, but rather with performing multiscale analysis (Levin, 1992).

A method employing conditional entropy profiles has been developed for obtaining a multiresolution characterization of the fragmentation pattern of a landscape within a fixed geographic extent (Johnson, Tempelman and Patil, 1995; Johnson and Patil, 1998). In addition, a parametric stochastic model (HMTM model) has been defined for simulating landscapes from known generating mechanisms so that the profile method can be evaluated under specified “null” scenarios (Johnson, Myers and Patil, 1999). In this paper, we present a method for computing conditional entropy profiles in terms of the parameters of the foregoing simulation model. These are called *process* profiles in contrast to the *empirical* profiles that are calculated from an actual map. Process profiles are presented corresponding to a range of model parameters, including some parameter values that were chosen to result in spatial patterns similar to three watershed-delineated landscapes in Pennsylvania.

2. Conditional entropy profiles for characterizing landscape fragmentation

We are interested in categorical raster maps with rectangular pixels and with K distinct mapping categories. For definiteness, we refer to the mapping categories as landcover types. Consider a sequence $\mathcal{G}_0, \mathcal{G}_1, \dots$ of such maps with the same spatial extent but of increasingly finer resolution. Specifically, we suppose that the map \mathcal{G}_0 consists of a single pixel and that each pixel of \mathcal{G}_n is bisected horizontally and vertically to produce the pixels of \mathcal{G}_{n+1} (Fig. 1). When attention is focused on two consecutive resolutions, \mathcal{G}_n and \mathcal{G}_{n+1} , the pixels of \mathcal{G}_n will be referred to as “parent” pixels and those of \mathcal{G}_{n+1} as “child” pixels. Thus, every parent has four children.

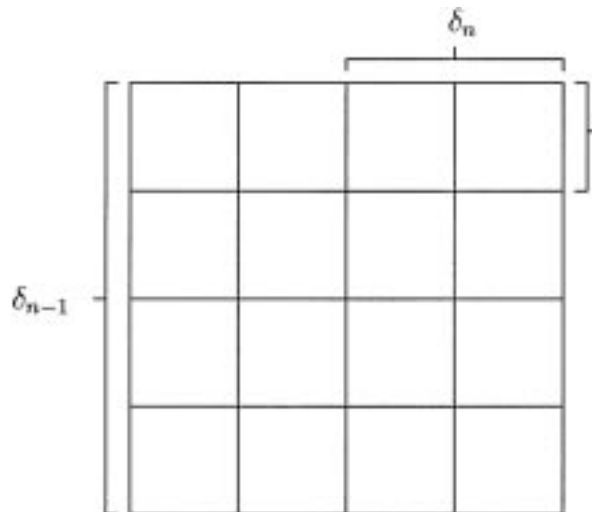


Figure 1. Hierarchical nesting of pixels.

The relative frequency of occurrence of landcover type i across \mathcal{G}_n will be denoted by $\widehat{P}_1^{[n]}, i = 1, \dots, K$. The ‘‘hat’’ indicates that these quantities are calculated from data. Thus, $\widehat{P}_1^{[n]} + \widehat{P}_2^{[n]} + \dots + \widehat{P}_K^{[n]} = 1$. Each pixel of \mathcal{G}_n has four children and the landcover types j_1, j_2, j_3, j_4 of these children comprise a 4-tuple denoted by $s = (j_1, j_2, j_3, j_4)$, where the ordering is according to the pixel sequence: NW corner, NE corner, SW corner, SE corner. The number of possible 4-tuples is K^4 . In a typical application, K might be 8 and $K^4 = 4096$. Comparing map \mathcal{G}_{n+1} with \mathcal{G}_n , we let $\widehat{P}_{i,s}^{[n,n+1]}$ be the relative frequency of occurrence of 4-tuple s across \mathcal{G}_{n+1} given that the parent cell in \mathcal{G}_n has landcover type i . Note that the array $\widehat{P}_{i,s}^{[n,n+1]}$ has $K \times K^4 = K^5$ entries. If $K = 8$, as above, this amounts to some 32,000 entries.

The conditional entropy of \mathcal{G}_{n+1} given \mathcal{G}_n is defined to be

$$\begin{aligned} \widehat{H}_n &\equiv \widehat{H}(\mathcal{G}_{n+1} | \mathcal{G}_n) \\ &= - \sum_{i=1}^K \widehat{P}_i^{[n]} \sum_s \widehat{P}_{i,s}^{[n,n+1]} \log_2 \widehat{P}_{i,s}^{[n,n+1]}, \end{aligned} \quad (1)$$

where $x \log_2 x$ is taken to equal 0 when $x = 0$. Conditional entropy satisfies the bounds,

$$0 \leq \widehat{H}_n \leq \log_2(K^4).$$

The lower bound is achieved when $\widehat{P}_{i,s}^{[n,n+1]} = 0$ for all i and for all but one 4-tuple s (which may depend on i). In other words, the landcover types of the four children are completely determined by the type of their parent. Similarly, the upper bound occurs when $\widehat{P}_{i,s}^{[n,n+1]} = 1/K^4$ for all i and all s . This means that the landcover types of the four children are assigned independently and at random (in the equal probability sense) regardless of the parental type. Following Colwell (1974), the upper bound implies zero predictability of the fine scale map \mathcal{G}_{n+1} given the coarse scale map \mathcal{G}_n . Conversely, the lower bound implies complete predictability of \mathcal{G}_{n+1} given \mathcal{G}_n .

The entropy \widehat{H}_n may be plotted against resolution to give a *conditional entropy profile* or *fragmentation profile*. In drawing the graph, the resolutions are arranged along the horizontal axis from finest resolution on the left to coarsest resolution on the right.

Two difficulties arise at this point. First, one seldom has multiresolution data on a hierarchically nested series of pixel grids. We address this problem by starting with a single, high resolution map and apply a suitable filter to generate the needed coarser resolution maps. This procedure is described in the next section.

The second difficulty is that the $\widehat{P}_{i,s}^{[n,n+1]}$ and the \widehat{H}_n are subject to severe fluctuation unless the maps are fairly large, thus limiting the resolutions for which the entropy profile can be reliably plotted. We can see the reason for this by supposing that \mathcal{G}_n is 512×512 and \mathcal{G}_{n+1} is 1024×1024 . This provides 512^2 parent pixels and correspondingly many 4-tuples of child pixels to compute the K^5 relative frequencies $\widehat{P}_{i,s}^{[n,n+1]}$. On average, then, there are $512^2/K^5$ observations per entry in $\widehat{P}_{i,s}^{[n,n+1]}$. When $K = 8$, this gives an average of 8 observations per entry, which is not too bad. But, as the resolutions are successively halved, the average number of observations per entry falls off as 8, 2, 1/2, 1/8, . . ., which is small enough to ensure that many of the $\widehat{P}_{i,s}^{[n,n+1]}$ will be zero. We get around this difficulty by replacing the $\widehat{P}_{i,s}^{[n,n+1]}$ by suitable expected values,

$$P_{i,s}^{[n,n+1]} = E[\widehat{P}_{i,s}^{[n,n+1]}],$$

where the expectations are taken with respect to the filter which is stochastic, as explained in the next section.

2.1 Obtaining coarser resolution maps with a random filter

For a given sensor type, image data is provided at one resolution, such as the 30-meter resolution of the Landsat thematic mapper. Therefore, given an actual fixed resolution data set in the form of a raster map of landcover types, coarser scale maps need to be produced by some protocol for aggregating the landcover types of child pixels into a landcover type for their parent. One class of such protocols, known as *resampling filters*, selects one of the child types for assignment to the parent pixel. Examples of resampling filters include (i) selecting the most frequent of the four child types (modal filter, Benson and MacKenzie, 1995), and (ii) selecting the landcover type at a fixed position (e.g., northwest corner) among the four child pixels (Costanza and Maxwell, 1994).

We have adopted a *random filter* which selects from the four child positions with equal probability. It can be regarded as an approximation to the modal filter. The fact that the coarser resolution maps are random is actually an advantage since it gives meaning to the expected values mentioned at the end of the previous section.

2.2 Computing empirical fragmentation profiles

The original datamap provides the ‘‘floor resolution’’ map \mathcal{G}_L which is $2^L \times 2^L$. In case the datamap is not square or has a ragged boundary, it is embedded in \mathcal{G}_L and pixels of \mathcal{G}_L that are outside the boundary of the datamap are masked with a special ‘‘nodata’’ code. The random filter is then applied to generate a sequence, $\mathcal{G}_L, \mathcal{G}_{L-1}, \dots, \mathcal{G}_0$, of generalized raster maps of increasingly coarser resolution. Specifically, \mathcal{G}_n is obtained by applying the random filter to 2×2 windows of \mathcal{G}_{n+1} . These windows do not overlap—they are window panes rather than sliding windows. The filter treats ‘‘nodata’’ as a $(K + 1)$ st category. Except for the floor resolution \mathcal{G}_L , the generalized maps are random due to randomness in the filter. We wish to average out this randomness in calculating entropy profiles.

It might be tempting to simulate the random filter many times, calculating conditional entropy for each replication and then averaging the $\hat{H}_n(\mathcal{G}_{n+1}|\mathcal{G}_n)$ values across replications. However, sample entropy is well-known to be a biased estimator of true entropy (Basharin, 1959), so that $E[\hat{H}_n(\mathcal{G}_{n+1}|\mathcal{G}_n)]$ is not what we are looking for. It is preferable to replace empirical relative frequencies by their expected values and calculate entropy on these expected relative frequencies. It turns out that the latter can be calculated theoretically without having to simulate the random filter.

2.2.1 Expected relative frequencies under the random filter

For consecutive resolutions, n and $n + 1$, consider the joint probability $\pi_{i,s}^{[n,n+1]}$ that a parent pixel in \mathcal{G}_n has landcover type i and the four child pixels in \mathcal{G}_{n+1} have their landcover types given by the 4-tuple $s = (j_1, j_2, j_3, j_4)$. These probabilities are with respect to the random filter. Marginal distributions of parent types and of 4-tuples of child types are denoted by

$$P_i^{[n]} = \sum_s \pi_{i,s}^{[n,n+1]}, \quad (2)$$

$$P_s^{[n+1]} = \sum_{i=1}^K \pi_{i,s}^{[n,n+1]}. \quad (3)$$

The conditional distribution of s given the parent type i is

$$P_{i,s}^{[n,n+1]} = \frac{\pi_{i,s}^{[n,n+1]}}{P_i^{[n]}}, \quad (4)$$

and the conditional distribution of the parent type given the 4-tuple of child types s becomes

$$P_{s,i}^{[n+1,n]} = \frac{\pi_{i,s}^{[n,n+1]}}{P_s^{[n+1]}}. \quad (5)$$

The probabilities $P_i^{[n]}$ and $P_{i,s}^{[n,n+1]}$ are the theoretical counterparts of the relative frequencies $\widehat{P}_i^{[n]}$ and $\widehat{P}_{i,s}^{[n,n+1]}$ discussed earlier. For purposes of obtaining the conditional entropy profile, $P_{i,s}^{[n,n+1]}$ is most important (see Equation (1)). Unfortunately, it is also the most difficult to compute directly. However, the four sets of probabilities in Equations (2), (3), (4), and (5) are related by Bayes formula so knowledge of any three sets determines the fourth set. The $P_{s,i}^{[n+1,n]}$ are the conditional probabilities of passing from fine resolution map to coarser resolution map and these probabilities are determined by the random filter. For example, if $s = (1, 4, 1, 2)$ then $P_{s,i}^{[n+1,n]}$ is $1/2$ when $i = 1$, is $1/4$ when $i = 2$ or $i = 4$, and is 0 for all other values of i . In general, if $s = (j_1, j_2, j_3, j_4)$ then

$$P_{s,i}^{[n+1,n]} = \frac{1}{4} (\delta_{j_1,i} + \delta_{j_2,i} + \delta_{j_3,i} + \delta_{j_4,i}), \quad (6)$$

where $\delta_{j,i}$ is the Kronecker delta which equals 1 when $j = i$ and equals 0, otherwise. Observe that the right hand side of Equation (6) does not depend on n or the original datamap.

At this point, it remains to calculate the two sets of marginal probabilities in Equations (2) and (3). Determination of $P_{i,s}^{[n,n+1]}$ requires a bit of work. It will be convenient to associate a color with each landcover type so that a categorical raster map assigns a color to each pixel. Likewise, a random map assigns a color *distribution* to each pixel. For a pixel u in \mathcal{G}_n let $f_u^{[n]}(\cdot)$ be its color distribution so that $f_u^{[n]}(i), i = 1, 2, \dots, K$, is the probability that pixel u has color i . Note that u is an aggregate of floor resolution pixels and that $f_u^{[n]}(i)$ is the relative frequency of color i among these pixels. If $v = (v_1, v_2, v_3, v_4)$ is the 4-tuple of child pixels of u (so the v_i are pixels in \mathcal{G}_{n+1}), the definition of the random filter implies that

$$f_u^{[n]}(i) = \frac{1}{4} (f_{v_1}^{[n+1]}(i) + f_{v_2}^{[n+1]}(i) + f_{v_3}^{[n+1]}(i) + f_{v_4}^{[n+1]}(i)). \quad (7)$$

This equation effectively replaces the random filter on color maps by a deterministic *linear filter* on spatially referenced color distributions.

The map \mathcal{G}_L at the floor resolution is deterministic so its color distribution is degenerate for each pixel: $f_v^{(L)}(i) = 1$ if pixel v has color i and $f_v^{(L)}(i) = 0$, otherwise. Formally,

$$f_v^{(L)}(i) = \delta_{A,i}, \quad (8)$$

where A is the color of pixel v . The color distributions at other resolutions are obtained by applying Equation (7) iteratively.

Let $\mathbf{v} = (v_1, v_2, v_3, v_4)$ be a 4-tuple of child pixels in \mathcal{G}_{n+1} and $\mathbf{s} = (j_1, j_2, j_3, j_4)$ a 4-tuple of landcover types (colors). The probability that the colors \mathbf{s} are assigned to the pixels \mathbf{v} is

$$f_v^{[n+1]}(\mathbf{s}) \equiv f_{v_1}^{[n+1]}(j_1) f_{v_2}^{[n+1]}(j_2) f_{v_3}^{[n+1]}(j_3) f_{v_4}^{[n+1]}(j_4). \quad (9)$$

This product formula is not true for random maps generally. But, with a little reflection, it is obviously true for deterministic maps, i.e., the floor resolution map. It is true for the other resolutions because the random filter is applied independently in different window panes. The marginal distribution $P_s^{[n+1]}$ in Equation (3) is obtained by averaging the expression (9) across the w_{n+1} window panes \mathbf{v} in \mathcal{G}_{n+1} to give

$$P_s^{[n+1]} = \frac{1}{w_{n+1}} \sum_{\mathbf{v}} f_{v_1}^{[n+1]}(j_1) f_{v_2}^{[n+1]}(j_2) f_{v_3}^{[n+1]}(j_3) f_{v_4}^{[n+1]}(j_4). \quad (10)$$

However, if masking of floor resolution pixels is necessary, we must delete 4-tuples \mathbf{s} containing ‘‘nodata’’ values and renormalize the $P_s^{[n+1]}$ table.

Next we consider the marginal probabilities $P_i^{[n]}$ in Equation (2). It may seem intuitive that these probabilities should not depend upon the resolution n . This is true (as will be proved shortly) but only when the datamap is square of size $2^L \times 2^L$. If pixel-masking is necessary, then $P_i^{[n]}$ can vary slightly with the resolution n . First, observe that

$$P_i^{[n]} = \frac{1}{N_n} \sum_u f_u^{[n]}(i), \quad (11)$$

where N_n is the number of pixels in \mathcal{G}_n and the sum is over all pixels u in \mathcal{G}_n . Again, deletion of the ‘‘nodata’’ value i and renormalization of the table is required in case of pixel-masking. Summing both sides of Equation (7) with respect to u and using the fact that $4N_n = N_{n+1}$ establishes the earlier claim that $P_i^{[n]}$ does not depend upon n when there is no pixel-masking. The argument breaks down when there is pixel-masking because the sum must exclude the ‘‘nodata’’ values.

2.2.2 Computer algorithm

A computer program has been written which uses Equations (7), (8), (11), and (10) to compute the quantities $P_i^{[n]}$ and $P_s^{[n+1]}$ for any given datamap. We briefly sketch the algorithm without going into coding details. A bank of four memory buffers is allocated for each of the maps $\mathcal{G}_L, \mathcal{G}_{L-1}, \dots, \mathcal{G}_1$. Each memory buffer contains K double precision locations to hold the color distribution for a single pixel. (In fact, $K + 1$ locations are used to account for the possibility of ‘‘nodata’’ pixels.) At most four color distributions per map are held in memory at any time, requiring only

$$4 \cdot (K + 1) \cdot L \cdot \text{size of (double)}$$

units of storage. Initially, all memory buffers are declared to be empty. Next, pixels of the floor resolution map (size $2^L \times 2^L$) are processed sequentially, but in quadtree order instead of row-major order (see Table 1). As each floor resolution pixel is processed, its degenerate

color distribution (Equation (8)) is stored in the first empty memory buffer for \mathcal{G}_L . Whenever all four memory buffers for a map (say, \mathcal{G}_{n+1}) become full, five things happen:

- (1) Formula (10) is applied to update a table storing the $P_s^{[n+1]}$ values. However, no updating is done for 4-tuples $s = (j_1, j_2, j_3, j_4)$ containing “nodata” values.
- (2) Formula (11), with $n + 1$ replacing n , is applied to all four memory buffers to update a table of $P_i^{[n+1]}$ values. Again, no updating is done for “nodata” values.
- (3) The right-hand side of Equation (7) is computed and stored in the next available memory buffer for \mathcal{G}_n . A flag is set if this fills the four memory buffers for \mathcal{G}_n .
- (4) All four memory buffers for \mathcal{G}_{n+1} are declared to be empty.
- (5) If the flag in step (3) is set, go to step (1) with \mathcal{G}_n as the active map. Otherwise, process the next floor resolution pixel.

2.2.3 Empirical profile calculation

After the computer program is run, the marginals $P_i^{[n]}$ and $P_s^{[n+1]}$ are available for profile calculation. As pointed out above, the marginal probability (4) can be determined by Bayes formula. But for programming ease and storage efficiency, it is better to use the entropy decomposition formula (Patil and Taillie, 1979), which is a consequence of Bayes law, and states that the total entropy of a two-way table is the entropy of either marginal plus the conditional entropy given that marginal. This gives

$$H(\text{parent}) + H(\mathcal{G}_{n+1}|\mathcal{G}_n) = H(\text{child}) + H(\mathcal{G}_n|\mathcal{G}_{n+1}),$$

so that the conditional entropy profile values are obtained as

$$H_n \equiv H(\mathcal{G}_{n+1}|\mathcal{G}_n) = H(\mathcal{G}_n|\mathcal{G}_{n+1}) + H(\text{child}) - H(\text{parent}), \quad (12)$$

where

$$H(\mathcal{G}_n|\mathcal{G}_{n+1}) = - \sum_s P_s^{[n+1]} \sum_{i=1}^K P_{s,i}^{[n+1,n]} \log_2(P_{s,i}^{[n+1,n]}), \quad (13)$$

Table 1. Sequence numbers for the quadtree ordering of pixels in an 8×8 lattice. The basic sequencing rule for quadtree ordering is (NW corner, NE corner, SW corner, SE corner), which is applied hierarchically until an entire $2^L \times 2^L$ array of pixels has been sequenced. There are simple rules relating sequence numbers for quadtree ordering to sequence numbers for row-major ordering.

0	1	4	5	16	17	20	21
2	3	6	7	18	19	22	23
8	9	12	13	24	25	28	29
10	11	14	15	26	27	30	31
32	33	36	37	48	49	52	53
34	35	38	39	50	51	54	55
40	41	44	45	56	57	60	61
42	43	46	47	58	59	62	63

$$H(\text{child}) = - \sum_s P_s^{[n+1]} \log_2(P_s^{[n+1]}), \quad (14)$$

$$H(\text{parent}) = - \sum_{i=1}^K P_i^{[n]} \log_2(P_i^{[n]}). \quad (15)$$

Equation (12) together with Equations (13), (6), (14), and (15) express the conditional entropy profile values H_n in terms of the marginals $P_i^{[n]}$ and $P_s^{[n+1]}$.

The formidable-looking expression,

$$Q_s = - \sum_{i=1}^K P_{s,i}^{[n+1,n]} \log_2(P_{s,i}^{[n+1,n]}), \quad (16)$$

which appears in Equation (13) is actually quite simple and does not depend on the resolution n or on the floor resolution map. We are going to show that Q_s takes on only five distinct values as s ranges over the K^4 different 4-tuples. First observe, from Equation (6), that Q_s is unchanged when the components of $s = (j_1, j_2, j_3, j_4)$ are permuted. The 4-tuples s can be classified into four different types depending on the number of distinct colors among j_1, j_2, j_3, j_4 . The 4-tuples of type 2 (i.e., those with exactly two distinct colors) can be further classified into two subtypes depending on whether the two colors are distributed among the four pixels in a 50–50 split or a $\frac{1}{4}$ – $\frac{3}{4}$ split. Table 2 lists the five different types of 4-tuples s and the corresponding values of Q_s . The second column of the table gives a canonical permutation of the components of s for each type.

Equation (13) now takes the simple form of a linear combination of the probabilities (with respect to the random filter) of occurrence of the different types of 4-tuples in \mathcal{G}_{n+1} :

$$\begin{aligned} H(\mathcal{G}_n | \mathcal{G}_{n+1}) &= \sum_s P_s^{[n+1]} Q_s \\ &= 0.81 \sum_{s:\text{type 2a}} P_s^{[n+1]} + \sum_{s:\text{type 2b}} P_s^{[n+1]} \\ &\quad + \frac{3}{2} \sum_{s:\text{type 3}} P_s^{[n+1]} + 2 \sum_{s:\text{type 4}} P_s^{[n+1]}. \end{aligned} \quad (17)$$

Combining Equations (12), (14), (15) and (17) gives an explicit expression,

Table 2. The five types of 4-tuples s and corresponding values of Q_s in Equation (16). The second column gives a canonical permutation of the components of s and the third column is obtained from Equation (6).

Type	$j_1 j_2 j_3 j_4$	$P_{s,i}^{[n+1,n]}$	Q_s
1	AAAA	δ_{Ai}	$-1 \log_2 1 = 0$
2a	AAAB	$\frac{3}{4} \delta_{Ai} + \frac{1}{4} \delta_{Bi}$	$-\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 0.81$
2a	AABB	$\frac{1}{2} \delta_{Ai} + \frac{1}{2} \delta_{Bi}$	$-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1$
3	AABC	$\frac{1}{2} \delta_{Ai} + \frac{1}{4} \delta_{Bi} + \frac{1}{4} \delta_{Ci}$	$-\frac{1}{2} \log_2 \frac{1}{2} - \frac{2}{4} \log_2 \frac{1}{4} = \frac{3}{2}$
4	ABCD	$\frac{1}{4} (\delta_{Ai} + \delta_{Bi} + \delta_{Ci} + \delta_{Di})$	$-\frac{4}{4} \log_2 \frac{1}{4} = 2$

$$\begin{aligned}
H_n = & - \sum_s P_s^{[n+1]} \log_2(P_s^{[n+1]}) + \sum_{i=1}^K P_i^{[n]} \log_2(P_i^{[n]}) \\
& + 0.81 \sum_{s:\text{type 2a}} P_s^{[n+1]} + \sum_{s:\text{type 2b}} P_s^{[n+1]} \\
& + \frac{3}{2} \sum_{s:\text{type 3}} P_s^{[n+1]} + 2 \sum_{s:\text{type 4}} P_s^{[n+1]}, \tag{18}
\end{aligned}$$

for the profile values in terms of $P_i^{[n]}$ and $P_s^{[n+1]}$. Our computer program calculates the latter two quantities as described in Section 2.2.2.

3. Process profiles for the HMTM landscape simulator

The method discussed in Section 2.2 is to be used for assessing multiscale spatial pattern of an actual landscape when only the floor resolution is available as a dataset. Prior to such analysis, it is desired to understand profile behavior for landscapes whose realization at the floor resolution arose from known generating processes.

A modeling approach that simulates categorical raster maps with hierarchically scaled spatial patterns that are similar to those of actual landscapes has been developed by Johnson, Myers and Patil (1999). This method defines a hierarchical landscape generating process using a series of Markov transition matrices. Starting from the coarsest scale with a single-pixel map, the assignment of one of K landcover categories to this pixel is determined by a probability vector $\mathbf{p}^{[0]} = [p_1^{[0]}, \dots, p_K^{[0]}]$. Subsequent transitions to the scales $1, \dots, L$ are determined by row-stochastic matrices whose elements, $G_{i,j}^{[n,n+1]}$, give the conditional probability of transition from parent pixel with landcover category i to a single child pixel with landcover category j . An important, but limiting, aspect of the model is that conditional on the parent pixel, landcover categories are assigned to its four child pixels in an *exchangeable and independent* manner. This model specification could be relaxed but the number of parameters in the model would become very large. Already, the model parameters comprise a K -dimensional initial vector $\mathbf{p}^{[0]}$ and a $K \times K$ transition matrix $\mathbf{G}^{[n,n+1]}$ for each transition. The simulation model is referred to as the *Hierarchical Markov Transition Matrix* (HMTM) model.

The multi-step transition matrices in the HMTM model are obtained by matrix multiplication,

$$\mathbf{G}^{[n_1, n_2]} = \mathbf{G}^{[n_1, n_1+1]} \mathbf{G}^{[n_1+1, n_1+2]} \dots \mathbf{G}^{[n_2-1, n_2]},$$

and the process landcover marginal, $\mathbf{p}^{[n]}$, after n transitions has its components given by

$$p_j^{[n]} = \sum_{i=0}^K p_i^{[0]} G_{i,j}^{[0,n]} \tag{19}$$

Running the simulation model for L transitions generates a $2^L \times 2^L$ floor resolution map \mathcal{G}_L to which the random filter can be applied. Each realization of \mathcal{G}_L thus determines a fragmentation profile, but we would like to somehow average across all the realizations to obtain a ‘‘process’’ profile. The formula (18) indicates how the entropy profile is

calculated in terms of the marginals $P_i^{[n]}$ and $P_s^{[n+1]}$. We will obtain the process profile by replacing these two sets of marginals by their averages across all possible realizations of the generating model but we would like to calculate these averages directly from the parameters of the model rather than estimate them by simulation. Under the HMTM model, responses in different pixels of \mathcal{G}_L are identically distributed with distribution given by the process marginal $\mathbf{p}^{[L]}$. Since $P_i^{[L]}$ is the average of the response indicators across all pixels in \mathcal{G}_L it follows that

$$E[P_i^{[L]}] = p_i^{[L]},$$

where the expectation is with respect to the process. Assuming that model output is not masked, we have

$$E[P_i^{[n]}] = p_i^{[L]}, \quad (20)$$

for all n since $P_i^{[n]} = P_i^{[L]}$ for all n . Because of Equation (19), this determines the expected $P_i^{[n]}$ in terms of the process parameters.

Using the characterization of the random filter in terms of Kronecker deltas (formula 6), it is tedious but straightforward to show that

$$E[P_s^{[n+1]}] = \sum_{i=1}^K P_i^{[n]} \cdot (G_{i,j_1}^{[n,L]} \cdot G_{i,j_2}^{[n,L]} \cdot G_{i,j_3}^{[n,L]} \cdot G_{i,j_4}^{[n,L]}), \quad (21)$$

where $s = (j_1, j_2, j_3, j_4)$. In fact, this equation can be obtained more directly as follows (on first reading, take $n+1$ equal to L , then to $L-1$, etc.). In order to obtain the map \mathcal{G}_{n+1} , the random filter has to be applied $t = L - n - 1$ times to \mathcal{G}_L . Each pixel in \mathcal{G}_{n+1} is an aggregate of a $2^t \times 2^t$ array of pixels in \mathcal{G}_L and, according to the random filter, its color is the color of a randomly selected pixel from this array. In computing $P_s^{[n+1]}$, we have to consider a 4-tuple of pixels from \mathcal{G}_{n+1} and a corresponding set of four arrays containing a total of $w = 4 \times 2^t \times 2^t = 2^{t+1} \times 2^{t+1}$ pixels from \mathcal{G}_L . Now consider these w pixels in relation to the simulation model. To find a common ancestor, we have to go back $t+1$ transitions to hierarchical level $L - (t+1) \equiv n$ in the model. Given the color i in this common ancestor, the responses in all w pixels are identically distributed with conditional probability of response j given by $G_{i,j}^{[n,L]}$. Moreover, the four sets of responses in the four arrays of \mathcal{G}_L are conditionally independent. (This is because the four children of the common ancestor are already conditionally independent.) It follows that

$$E[P_s^{[n+1]}|i] = G_{i,j_1}^{[n,L]} \cdot G_{i,j_2}^{[n,L]} \cdot G_{i,j_3}^{[n,L]} \cdot G_{i,j_4}^{[n,L]}. \quad (22)$$

Equation (21) follows by taking expectations with respect to i at level n in the hierarchy.

3.1 Process profiles for HMTM models chosen to simulate landscapes similar to Pennsylvania watersheds

The HMTM simulation model is fully specified by its sequence of transition matrices $\mathbf{G}^{[n,n+1]}$ and by the initial distribution $\mathbf{p}^{[0]}$. To simulate the null scenario of a self-similar fragmentation process, we take all of the transition matrices equal (to, say, \mathbf{G}). We also suppose that the initial vector is a stationary vector for \mathbf{G} . This constant sequence of

transition matrices could be altered at any stage in the hierarchy to induce breaks in self-similarity so that we may evaluate the sensitivity of a conditional entropy profile to detecting such breakpoints and the characteristic scaling domains of self-similarity between breakpoints. However, only self-similar HMTM models are considered in this paper.

Transition matrices have been obtained subjectively by Johnson, Myers and Patil (1999) to replicate landscape features of watersheds in Pennsylvania as mapped by eight landcover categories at a floor resolution of 30-meter pixels. (Formal methods for fitting the HMTM model to landcover maps will be treated in a forthcoming publication.) The landcover maps were obtained by classifying 6-band Landsat thematic mapper scenes. The matrices for three selected watersheds (Fig. 2) are reproduced here in Table 3. The Sinnemahoning Creek watershed is mostly forested, representing a continuum of forest interior wildlife habitat. The Jordan Creek watershed represents a transitional landscape that barely maintains a connected forest matrix which is encroached by agriculture and urban/suburban land use. Finally, the Conestoga Creek watershed (near Philadelphia) represents a landscape that is dominated by open agricultural land and highly aggregated urban/suburban land, with little remaining forest.

For each transition matrix in Table 3, Equations (20) and (21) were applied to obtain the two sets of expected marginals $E[P_i^{[n]}]$ and $E[P_s^{[n+1]}]$ for the randomly filtered series of maps. The conditional entropy profile for the process was then calculated from Equation (18). The resulting three fragmentation profiles are plotted in Fig. 3.

Parameters of the transition matrix model can be varied in many ways, thus allowing extensive experimentation to assess profile sensitivity to different spatial patterns. For example, the effect of diagonal dominance in the transition matrix, which should be associated with strong spatial dependence and large patches, can be studied in a controlled manner. As a simple example, let the diagonal entries have a common value λ and let the remaining probability $(1 - \lambda)$ be evenly distributed across the other entries in each row of the transition matrix. Process profiles for a selection of such models are presented in Fig. 4.

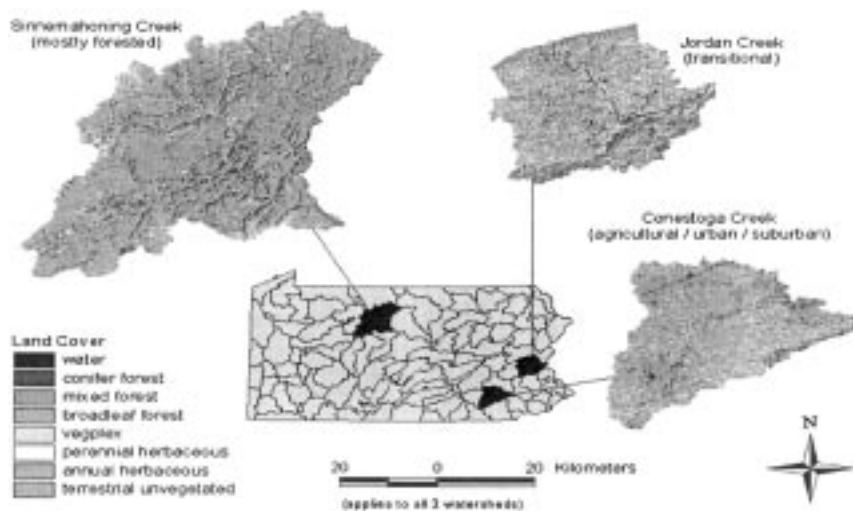


Figure 2. Landcover maps for three watersheds of Pennsylvania.

Table 3. Transition matrices modeling the landcover maps for the three watersheds in Fig. 2. There are $K = 8$ landcover categories. The matrices were obtained through subjective trial and error rather than any formal fitting procedure (Johnson, Myers and Patil, 1999).

<i>Mostly Forested, Based on Sinnemahoning Creek Watershed</i>								
	<i>W</i>	<i>C</i>	<i>M</i>	<i>B</i>	<i>VP</i>	<i>PH</i>	<i>AH</i>	<i>TU</i>
<i>W</i>	0.45	0.05	0.10	0.35	0.017	0.017	0.017	0
<i>C</i>	0.013	0.50	0.10	0.35	0.013	0.013	0.013	0
<i>M</i>	0.013	0.05	0.55	0.35	0.013	0.013	0.013	0
<i>B</i>	0.013	0.05	0.10	0.80	0.013	0.013	0.013	0
<i>VP</i>	0.017	0.05	0.10	0.35	0.45	0.017	0.017	0
<i>PH</i>	0.017	0.05	0.10	0.35	0.017	0.45	0.017	0
<i>AH</i>	0.017	0.05	0.10	0.35	0.017	0.017	0.45	0
<i>TU</i>	0.013	0.05	0.10	0.35	0.013	0.013	0.013	0.45
<i>Transitional, based on Jordan Creek Watershed</i>								
	<i>W</i>	<i>C</i>	<i>M</i>	<i>B</i>	<i>VP</i>	<i>PH</i>	<i>AH</i>	<i>TU</i>
<i>W</i>	0.495	0.050	0.100	0.125	0.100	0.050	0.08	0.00
<i>C</i>	0.010	0.525	0.100	0.125	0.100	0.050	0.08	0.01
<i>M</i>	0.010	0.050	0.575	0.125	0.100	0.050	0.08	0.01
<i>B</i>	0.010	0.050	0.100	0.600	0.100	0.050	0.08	0.01
<i>VP</i>	0.010	0.050	0.100	0.125	0.575	0.050	0.08	0.01
<i>PH</i>	0.010	0.050	0.100	0.125	0.100	0.525	0.08	0.01
<i>AH</i>	0.010	0.010	0.100	0.120	0.100	0.050	0.60	0.01
<i>TU</i>	0.000	0.000	0.000	0.075	0.100	0.075	0.00	0.75
<i>Agricultural/Urban/Suburban Mosaic based on Conestoga Creek Watershed</i>								
	<i>W</i>	<i>C</i>	<i>M</i>	<i>B</i>	<i>VP</i>	<i>PH</i>	<i>AH</i>	<i>TU</i>
<i>W</i>	0.550	0.062	0.062	0.100	0.062	0.062	0.100	0.00
<i>C</i>	0.010	0.565	0.070	0.100	0.070	0.070	0.105	0.01
<i>M</i>	0.010	0.070	0.565	0.100	0.070	0.070	0.105	0.01
<i>B</i>	0.010	0.040	0.040	0.700	0.100	0.050	0.050	0.01
<i>VP</i>	0.010	0.010	0.010	0.100	0.600	0.105	0.155	0.01
<i>PH</i>	0.010	0.010	0.085	0.100	0.105	0.600	0.080	0.01
<i>AH</i>	0.010	0.000	0.000	0.100	0.028	0.028	0.825	0.01
<i>TU</i>	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.90

W = water, C = conifer forest, M = mixed forest, B = broadleaf forest, VP = vegplex, PH = perennial herbaceous, AH = annual herbaceous, TU = terrestrial unvegetated.

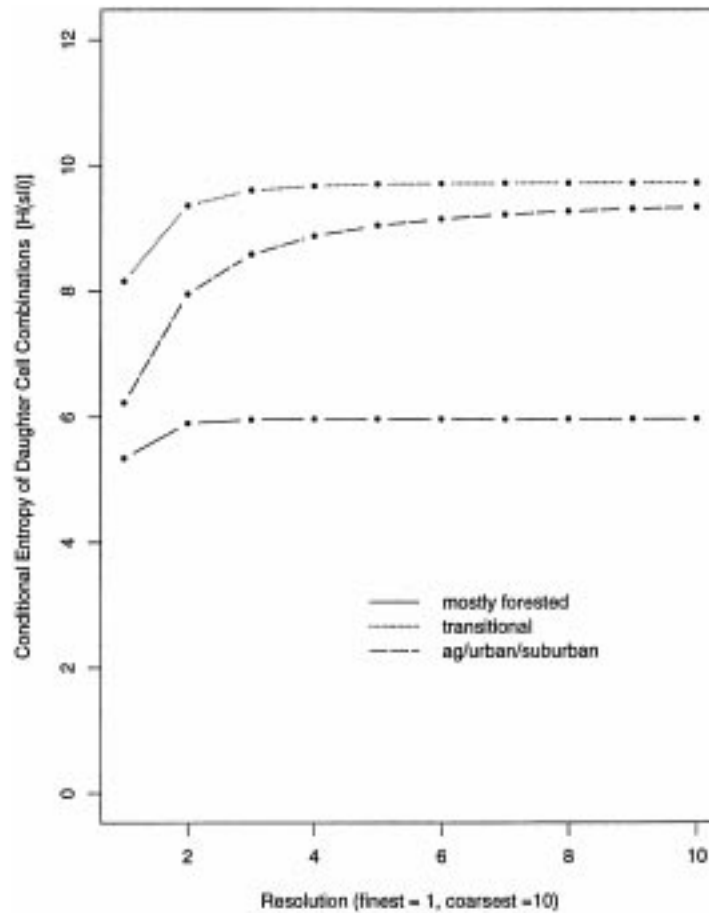


Figure 3. Conditional entropy process profiles for HMTM models whose transition matrices (Table 3) are obtained from watersheds with three distinctly different landcover patterns. The initial vector $P^{[0]}$ in the model was taken to be uniform across the $K = 8$ categories. It can be argued that the stationary vector for the transition matrix would be a better choice as initial vector.

4. Discussion

The results in Figs 3 and 4 reveal certain characteristics that may be expected from different landscape patterns. As particular landcover categories increase in both dominance and patch coherence, the conditional entropy at the floor resolution decreases. As map resolution is made coarser from a resampling filter, the resulting loss of information is revealed by increasing conditional entropy; however, the rate at which entropy increases as resolution decreases depends on the degree of patch coherence for each of the landcover categories. Patch coherence is regulated in the model by the degree of diagonal dominance of the transition matrix while marginal landcover dominance for any category is indicated by the corresponding column sum of the transition matrix.

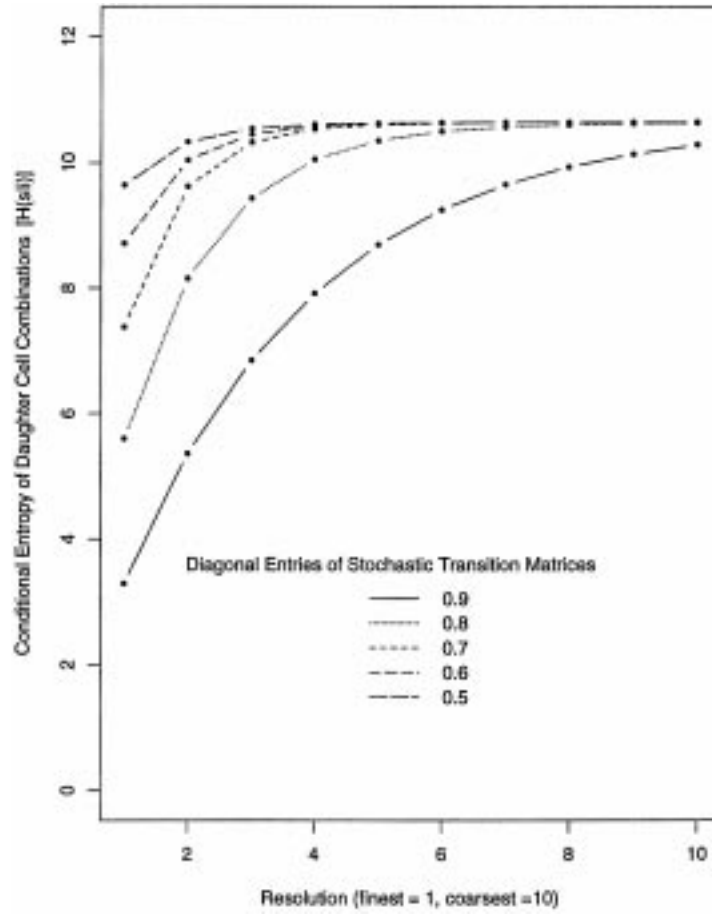


Figure 4. Conditional entropy process profiles for HMTM models whose hypothetical $K \times K$ transition matrices have the value λ along the diagonal and the value $(1 - \lambda)/(K - 1)$ off the diagonal. Here $K = 8$ and the values of λ are indicated in the legend. The stationary vector is uniform across the K categories and is used as the initial vector in the model. The floor resolution maps \mathcal{G}_L were 1024×1024 (i.e., $L = 10$). Large values of λ result in strong spatial dependence (indicated by large profile relief) which persists at larger distances (indicated by a slowly rising profile). All models have the same stationary vector and therefore the same horizontal asymptote.

Furthermore, as the number of landcover categories with only fine-grained patches increases, the maximum conditional entropy is achieved sooner as the resolution decreases. This is attributed to the filtering out of fine-grained patches as the map resolution is degraded by the filter. In this context, “fine-grained” means small patch sizes.

For the models based on characteristics of actual watershed-delineated landscapes in Pennsylvania, the mostly forested landscape is dominated by a single connected patch of broadleaf forest, with a sub-dominance of mixed and conifer forest, while other categories

are sparsely distributed throughout the landscape as fine-grained patches. Water is a relatively fine-grained landscape element in Pennsylvania; the state contains numerous rivers and streams but its mostly non-glaciated terrain has very few large lakes. Therefore, in a mostly forested landscape like the Sinnemahoning Creek watershed, even the floor resolution of 30-meter pixels is not fine-grained enough to capture much of the surface water. Accordingly, the conditional entropy profile flattens out fairly rapidly as the resolution coarsens (see lowest profile in Fig. 3). Maps with a flat or nearly flat profile behave like a random map in which the spatial distribution of landcover shows no particular structure beyond that which is imposed by dominance in the marginal landcover distribution.

The model based on the transitional landscape reveals a more even distribution of landcover categories; therefore, the maximum conditional entropy is highest because categories are retained as the resolution is degraded. Also, since no category exhibits both high patch coherence and high dominance over other categories, the conditional entropy at the floor resolution is high.

The model based on the mostly deforested landscape has a conditional entropy profile characteristic of a landscape with high patch coherence *and* co-dominance by several patch types, in this case annual herbaceous (mostly row crops), terrestrial unvegetated (mostly urban) and vegplex (composite vegetation types characteristic of suburbs and old fields).

Acknowledgments

Prepared with partial support from the United States Environmental Protection Agency (Environmental Statistics and Information Division, Office of Policy, Planning and Evaluation; and Environmental Monitoring and Assessment Program, EMAP Design and Statistics Group) and the National Science Foundation (NSF/EPA Water and Watersheds Program). The contents have not been subjected to reviews of these agencies and therefore do not necessarily reflect the views of these agencies and no official endorsement should be inferred.

References

- Basharin, G.P. (1959) On a statistical estimate for the entropy of a sequence of independent random variables. *Theory of Probability and its Applications*, **4**, 333–36.
- Benson, B.J. and MacKenzie, M.D. (1995) Effects of sensor spatial resolution on landscape structure parameters. *Landscape Ecology*, **10**, 113–20.
- Colwell, R.K. (1974) Predictability, constancy, and contingency of periodic phenomena. *Ecology*, **55**, 1148–253.
- Costanza, R. and Maxwell, T. (1994) Resolution and predictability: An approach to the scaling problem. *Landscape Ecology*, **9**, 47–57.
- Johnson, G.D., Myers, W.L., and Patil, G.P. (1999) Stochastic generating models for simulating hierarchically structured multi-cover landscapes. *Landscape Ecology*, **14**, 413–521.
- Johnson, G.D. and Patil, G.P. (1998) Quantitative multiresolution characterization of landscape patterns for assessing the status of ecosystem health in watershed management areas. *Ecosystem Health*, **4**, 177–287.

- Johnson, G.D., Tempelman, A.K., and Patil, G.P. (1995) Fractal based methods in ecology: a review for analysis at multiple spatial scales. *Coenoses*, **10**, 123–231.
- Kotliar, N.B. and Wiens, J.A. (1990) Multiple scales of patchiness and patch structure: A hierarchical framework for the study of heterogeneity. *Oikos*, **59**, 253–360.
- Levin, S. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–2067.
- O'Neill, R.V., Gardner, R.H., and Turner, M.G. (1992) A hierarchical neutral model for landscape analysis. *Landscape Ecology*, **7**, 55–61.
- O'Neill, R.V., Johnson, A.R., and King, A.W. (1989) A hierarchical framework for the analysis of scale. *Landscape Ecology*, **3**, 193–205.
- Patil, G.P. and Taillie, C.T. (1979) An overview of diversity. In *Statistical Ecology, Vol. 6, Ecological Diversity in Theory and Practice*, Grassle, J.F., Patil, G.P., Smith, W., and Taillie, C. (eds), International Co-operative Publishing House, Fairland, Maryland.
- Wiens, J.A. (1995) Landscape mosaics and ecological theory. In *Mosaic Landscapes and Ecological Processes*, Hansson, L., Fahrig, L., and Merriam, G. (eds), Chapman and Hall, London, pp. 1–26.

Biographical sketches

Glen Johnson has developed a multidisciplinary career in environmental science, whose focus has ranged from small scale issues of chemical toxicology and microbiology to large scale issues of geography, landscape ecology and remote sensing. Glen has a B.S. in Biology from the State University of New York College of Environmental Science and Forestry, along with a M.S. in Ecology, a M.A. in Statistics and a Ph.D. (1999) in Quantitative Ecology, all from Penn State University. His applied endeavors include six years with the Pennsylvania Department of Environmental Protection, and private consulting in Philadelphia. Currently, Glen works for the New York State Department of Health where he specializes in geographic and environmental epidemiology. His focus is on the incorporation of spatial autocorrelation into generalized linear models and the matching of data obtained at disparate spatial resolutions.

Wayne Myers is Associate Professor of Forest Biometrics in the School of Forest Resources, and Codirector of the Office for Remote Sensing of Earth Resources in the Environmental Resources Research Institute at the Pennsylvania State University. His research interests include remote sensing, geographic information systems, landscape ecology, artificial intelligence, and forest resources inventory.

Ganapati P. Patil is Distinguished Professor of Mathematical Statistics and Director of the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at the Pennsylvania State University.

C. Taillie is Senior Research Associate in the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at the Pennsylvania State University.