Landscape Ecology **16:** 597–610, 2001.

Research Article

Characterizing watershed-delineated landscapes in Pennsylvania using conditional entropy profiles

Glen D. Johnson[1,3,4,*], Wayne L. Myers[2,3], Ganapati P. Patil[1,3] & Charles Taillie[1]
[1]*Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Penn State University, 421 Thomas Building, University Park, PA 16802 USA;*[2]*School of Forest Resources and Environmental Resources Research Institute, Penn State University, 124 Land and Waters Building, University Park, PA 16802 USA;*[3]*Intercollege Graduate Degree Program in Ecology, Penn State University, 208 Mueller Laboratory, University Park, PA 16802 USA;*[4]*Current address: Stantec Consulting, Inc., 950 Industrial Way.* **(author for correspondence; e-mail: kjohnson@stantec.com)*

Received 15 November 1999; Revised 18 June 2001; Accepted 15 June 2001

Key words: conditional entropy profiles, landscape comparisons, landscape ecology, landscape monitoring, landscape pattern measurements, watersheds

Abstract

When the objective is to characterize landscapes with respect to relative degree and type of forest (or other critical habitat) fragmentation, it is difficult to decide which variables to measure and what type of discriminatory analysis to apply. It is also desirable to incorporate multiple measurement scales. In response, a new method has been developed that responds to changes in both the marginal and spatial distributions of land cover in a raster map. Multiscale features of the map are captured in a sequence of successively coarsened resolutions based on the random filter for degrading raster map resolutions. Basically, the entropy of spatial pattern associated with a particular pixel resolution is calculated, conditional on the pattern of the next coarser 'parent' resolution. When the entropy is plotted as a function of changing resolution, we obtain a simple two-dimensional graph called a 'conditional entropy profile', thus providing a graphical visualization of multi-scale fragmentation patterns.

Using eight-category raster maps derived from 30-meter resolution LANDSAT Thematic Mapper images, the conditional entropy profile was obtained for each of 102 watersheds covering the state of Pennsylvania (USA). A suite of more conventional single-resolution landscape measurements was also obtained for each watershed using the FRAGSTATS program. After dividing the watersheds into three major physiographic provinces, cluster analysis was performed within each province using various combinations of the FRAGSTATS variables, land cover proportions and variables describing the conditional entropy profiles. Measurements of both spatial pattern and marginal land cover proportions were necessary to clearly discriminate the watersheds into distinct clusters for most of the state; however, the Piedmont province essentially only required the land cover proportions. In addition to land cover proportions, only the variables describing a conditional entropy profile appeared to be necessary for the Ridge and Valley province, whereas only the FRAGSTATS variables appeared to be necessary for the Appalachian Plateaus province. Meanwhile, the graphical representation of conditional entropy profiles provided a visualization of multi-scale fragmentation that was quite sensitive to changing pattern.

Introduction

Assessment and monitoring of landscape-scale ecosystems requires quantitative characterization of spatial land cover patterns. This has resulted in the development of a plethora of measurements, many of which have been very conveniently programmed and made freely available by the U.S. Forest Service (McGarigal and Marks 1995). These measurements are quite valuable in their own right for obtaining information about specific landscape variables; however, one can rapidly get confused about what are the appropriate variables to be measuring when the objective is to categorize different landscapes into common groups. Such an objective may arise when many landscape units need to be assessed over a large region for the purpose of ranking units requiring intervention or remediation with respect to environmental degradation (Jones et al. 1997). Similarly, for any landscape unit that is being monitored over time, one may desire to categorize the status of the unit to determine if and how it may be changing. Therefore, one must decide what set of variables are appropriate to measure to discriminate among different landscape types.

A further complication is to decide on an appropriate measurement resolution (or grain) and an appropriate spatial extent of a landscape unit. Both of these aspects of spatial scale are known to affect observed patterns in landscapes (Wiens 1989). When a landscape is represented by a land cover raster map, measurement scale is the resolution (size) of a data pixel, which is fixed by the imaging sensor. Resolution is well known to affect observed patterns and the ability to detect change in patterns in image-based raster maps (e.g., see Townshend and Justice 1988; Qi and Wu 1996).

Instead of letting resolution be a problem, an alternative approach is to embrace the effect of changing resolution as something that is much more informative about spatial pattern than measurements based on a fixed resolution. This approach motivated several researchers to investigate the application of fractal-based information theoretic measurements, as summarized by Johnson *et al.* (1995).

While the concept of a 'conditional entropy profile' will be detailed in the next section, the basic product is a graph of raster map entropy at a particular pixel resolution, conditional on the next coarser resolution, which is plotted as a function of increasing pixel size. Thus, conditional entropy is traced out, revealing a multi-resolution characterization of the fragmentation pattern of a landscape within a fixed geographic extent (Johnson and Patil 1998; Johnson *et al.* 1999). Such a graph is monotonic non-decreasing, reaching a 'point of no further change' in a manner similar to the more familiar variogram models. However, variograms apply to numerical data arising from point measurements, whereas conditional entropy profiles apply to categorical data (land cover types) that arise from a spatially synoptic raster image. Our intention is to present a more holistic, general assessment of pattern that is expected to encompass many aspects of pattern that would otherwise require a suite of different measurements.

The most closely related conventional measurement is contagion; however, conditional entropy quantifies the spatial pattern at different resolutions in a way that captures the pattern's probabilistic dependence on the pattern of the next coarser resolution. This approach is more in line with hierarchy theory (O'Neill *et al.* 1989) which suggests that patterns observed at a given resolution will constrain the patterns observed at finer resolutions and will in turn be constrained by coarser resolution patterns. Indeed, hierarchically-nested spatial patterns have been observed in actual landscapes (Kotliar and Wiens 1990; O'Neill *et al.* 1992). Furthermore, Frohn (1998, pp. 69 and 75) showed that when contagion was measured over a range of resolutions that were obtained from an 'increasing size modal filter,' the results were an unpredictable trend in contagion as a function of decreasing pixel size. This indicates that contagion may not be a viable candidate for obtaining a multi-resolution characterization of spatial pattern.

The behavior of conditional entropy profiles has been evaluated using simulated multi-cover landscapes that were created by stochastic landscape generating models (Johnson, Myers and Patil 1999). Results indicated that the profiles can be expected to yield distinct properties in response to different landscape types that were created by known fragmentation processes.

This manuscript presents the first actual application of conditional entropy profiles, which were obtained for the landscapes of 102 watersheds that constitute a complete coverage of the state of Pennsylvania (Figure 1).

The ultimate objective is to categorize the watersheds into common landscape types with respect to degree and type of forest fragmentation. First, it is desired to see if the addition of profile variables can improve the ability of single-resolution variables to sensibly categorize the watersheds into different landscape types; secondly, it is desired to see if conditional entropy profiles alone can do a reasonable job of discriminating among differing landscape types.

Data

The data used for all landscape measurements came from eight-category raster maps derived from 30-meter resolution LANDSAT Thematic Mapper images. Details of how the raw satellite data were processed to derive the raster maps is available through metadata located at the Pennsylvania Spatial Data Access web page (*http://www.pasda.psu.edu*), under the category of 'Terrabyte images'. The method and soft-

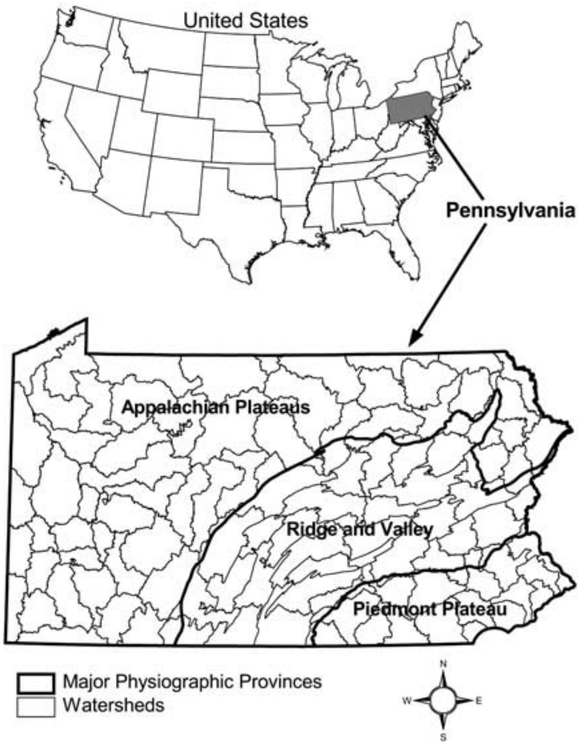

Figure 1. Location of the study area, Pennsylvania, within the United States (above) and boundaries of watersheds and major physiographic provinces within Pennsylvania (below).

ware is available as C-language programs for general use under the acronym PHASES (Myers 1999).

The land cover categories are *water, conifer forest, mixed forest, broadleaf forest, transitional, perennial herbaceous, annual herbaceous* and *terrestrial unvegetated*. The transitional category derives from a heterogeneous mix of land cover types; perennial herbaceous is primarily grassland that occurs in small patches, but occurs in larger patches where pasture land is present; annual herbaceous is primarily cropland and is often adjacent to patches of perennial herbaceous land; meanwhile, terrestrial unvegetated is primarily urbanized land. The remaining category labels are self explanatory.

Watershed delineations were obtained from the Penn State Environmental Resources Research Institute (ERRI), where small watersheds were aggregated to generally correspond to larger Pennsylvania State Water Plan watersheds. The source small watersheds were originally delineated by the Water Resources Division of the U.S. Geological Survey on 7.5-min topographic maps, then digitized to produce 9895 digital drainage basin boundaries in Pennsylvania.

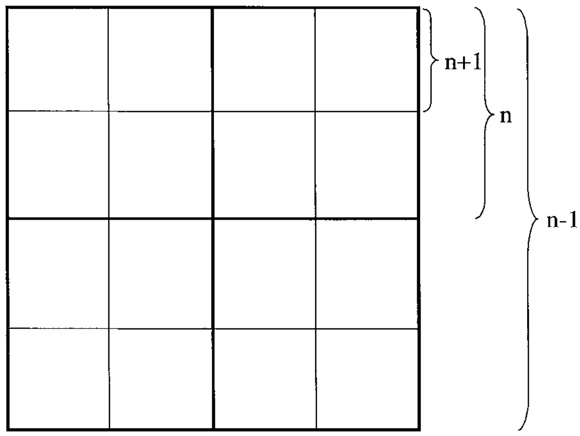

Figure 2. Hierarchical nesting of data pixels.

Measurements

Conditional entropy profiles

In order to conceptualize conditional entropy, consider a raster map whose pixels are each assigned one of K distinct land cover categories that are in turn represented by distinct colors. Now consider a map of the same extent that consists of four times as many data pixels that are hierarchically nested within the larger pixels of the first map. This sequence of increasingly finer resolutions may be continued so that each pixel of a given nth scale (resolution) is sub-divided into four 'child' pixels at the $(n+1)$st scale, and is itself a child pixel of a 'parent' pixel from the $(n-1)$st scale. This process is portrayed in Figure 2.

Letting the coarsest scale map be level 0, and the finest scale (floor resolution) be level L, we thus have a sequence $M_0, \cdots, M_L$ of maps with increasingly finer resolution.

For the nth resolution, let $\hat{P}_i^{[n]}$ equal the proportion of parent pixels from the nth scaled map that are labeled as category i for $i = 1, \cdots, K$, where the 'hat' symbolizes a quantity that is calculated from data. For each set of four child pixels at scale $n+1$ that are nested within a parent pixel, let **s** be the vector containing a unique ordering of their four land cover categories. For each of the K^4 possible vectors **s**, let $\hat{P}_{\mathbf{s}}^{[n+1]}$ equal the proportion of child '4-tuples' that yield the vector **s** at scale $n+1$. Now define $\hat{P}_{i\mathbf{s}}^{[n,n+1]}$ as the proportion of '4-tuples' at scale $n+1$ that are of vector **s**, given that their parent pixel is of category i. Finally, define $\hat{P}_{\mathbf{s}i}^{[n+1,n]}$ as the proportion of parent pixels in category i, given that the child 4-tuple is of vector **s**.

For a particular scale n, the marginal entropy of the parent pixels from the n scaled map is:

$$\hat{H}^{[n]}_{\text{pixel}} = -\sum_{i=1}^{K} \hat{P}^{[n]}_i \log \hat{P}^{[n]}_i$$

This is the entropy of the marginal land cover distribution at scaling level n. It does not depend on the spatial arrangement of pixel categories at this scaling level. The marginal entropy of the child 4-tuples from the $n+1$ scaled map is:

$$\hat{H}^{[n+1]}_{\text{4-tuple}} = -\sum_{j=1}^{K^4} \hat{P}^{[n+1]}_{\mathbf{s}} \log \hat{P}^{[n+1]}_{\mathbf{s}}$$

This entropy does depend on the spatial arrangement of pixel categories. The conditional entropy of the parent scale categories, given the child scale 4-tuples is:

$$\hat{H}^{[n+1,n]}_{\text{pixel|4-tuple}} = -\sum_{j=1}^{K^4} \hat{P}^{[n+1]}_{\mathbf{s}} \sum_{i=1}^{K} \hat{P}^{[n+1,n]}_{\mathbf{s}i} \log \hat{P}^{[n+1,n]}_{\mathbf{s}i}$$

Finally, the conditional entropy of the child scale 4-tuples, given the parent scale categories is:

$$\hat{H}^{[n,n+1]}_{\text{4-tuple|pixel}} = -\sum_{i=1}^{K} \hat{P}^{[n]}_i \sum_{j=1}^{K^4} \hat{P}^{[n,n+1]}_{i\mathbf{s}} \log \hat{P}^{[n,n+1]}_{i\mathbf{s}}$$

All of these expressions use the convention that $x \log x = 0$ when $x = 0$.

Conditional entropy profiles are graphs of the last expression, $\hat{H}^{[n,n+1]}_{\text{4-tuple|pixel}}$, plotted as a function of resolution, n; however, this quantity is difficult to compute directly. Since the total entropy of cross-classified factors can be decomposed into among and within sources (e.g., Patil and Taillie 1982; Pielou 1975; Colwell 1974), the four entropy components defined above are related as:

$$\hat{H}^{[n]}_{\text{pixel}} + \hat{H}^{[n,n+1]}_{\text{4-tuple|pixel}} = \hat{H}^{[n+1]}_{\text{4-tuple}} + \hat{H}^{[n+1,n]}_{\text{pixel|4-tuple}}. \quad (1)$$

Therefore, if the other three entropy components are obtained, then $\hat{H}^{[n,n+1]}_{\text{4-tuple|pixel}}$ is readily solved for.

Upper and lower bounds on the conditional entropy are given by:

$$0 \le \hat{H}^{[n,n+1]}_{\text{4-tuple|pixel}} \le \log(K^4)$$

The lower bound is achieved when each conditional distribution of 4-tuples, given the parent pixel, is degenerate, meaning that for each i with $\hat{P}^{[n]}_i > 0$, there is a 4-tuple of land cover categories $\mathbf{s'} = \mathbf{s'}(i)$ so that $\hat{P}^{[n,n+1]}_{i\mathbf{s}'} = 1$ and $\hat{P}^{[n,n+1]}_{i\mathbf{s}} = 0$ for all $\mathbf{s} \neq \mathbf{s}'$. The upper bound is achieved when each conditional distribution is uniform, meaning that for each i with $\hat{P}^{[n]}_i > 0$, $\hat{P}^{[n,n+1]}_{i\mathbf{s}} = 1/K^4$ for all $\mathbf{s}$. In other words, the lower bound is achieved when the 4-tuple of child categories is uniquely determined by the category of the parent pixel, and the upper bound is achieved when there is an even distribution of categories among children 4-tuples that are nested within parent pixels of a common parent category. To use the terminology of Colwell (1974), the upper bound implies a state of zero predictability of the $n+1$ scale map given the n scale map, while the lower bound implies complete predictability of the $n+1$ scale map given the n scale map.

Conditional entropy of expected frequencies

Data for a true multi-resolution framework as described above is not available in practice. Therefore, we must start with an actual data set that provides the finest resolution raster map, then obtain a sequence of increasingly coarser maps by successive application of a resampling filter whereby parent pixels are assigned colors according to some function of their four child pixels. That function may be deterministic or it may be random. For example, Costanza and Maxwell (1994) have taken the parent category to be that of the northwest corner child pixel. This is a deterministic rule, but the resulting sequence $M_L, M_{L-1}, \cdots, M_0$ of maps and subsequent calculations depend upon the arbitrary choice of sampling corner. A modal filter (i.e. Benson and MacKenzie 1995), in which the parent category is chosen as the most frequent among the four child categories, is attractive but some rule must be applied to resolve ties.

Our approach uses a random filter that closely approximates the modal filter. For each parent pixel, one of its four child pixels is selected at random and the parent is assigned the category of the selected child. If one were to implement this filter with a random number generator, the sequence of filtered maps would depend on the starting seed for the generator and it would be necessary to somehow average across many replications of the filtering mechanism. Instead, we have obtained analytic expressions, and corresponding computational algorithms, to obtain the *expected* relative frequencies $P^{[n]}_i = E[\hat{P}^{[n]}_i]$, $P^{[n+1]}_{\mathbf{s}} = E[\hat{P}^{[n+1]}_{\mathbf{s}}]$ and $P^{[n+1,n]}_{\mathbf{s},i} = E[\hat{P}^{[n+1,n]}_{\mathbf{s},i}]$. Here, expectations are with respect to the random filter applied to a fixed floor resolution map M_L. See Johnson *et al.* (1998) for computational details and supporting theory.

A noteworthy feature of the random filter is that the expected marginal land cover proportions are constant across resolution, whereby for each category i $\hat{P}_i^{[L]} \equiv P_i^{[L]} = P_i^{[L-1]} = P_i^{[L-2]} = \cdots$. In fact, this is analytically true only for square images; some variation across resolution occurs when the image has ragged boundaries, although the effect is usually very small. For the 102 Pennsylvania watersheds, the above string of equalities was correct to at least three decimal places.

The formulas for conditional entropy profiles given previously may be applied using these expected proportions, $P_i^{[n]}$, $P_{\mathbf{s}}^{[n+1]}$ and $P_{\mathbf{s}i}^{[n+1,n]}$, instead of their empirical counterparts. This is the approach taken in the rest of the paper. A desirable side benefit is that we avoid the use of empirical proportions in entropy calculations, which are known to give biased estimates of entropy.

When conditional entropy is plotted as a function of decreasing resolution (increasing pixel size), the result is a monotonic non-decreasing curve that reflects the *rate and extent of information loss* as one degrades the raster map by coarsening the resolution. For a completely random map, in which all land cover categories have the same frequency and pixels are assigned categories at random, the curve is a horizontal line whose height is the maximum obtainable conditional entropy, $\log(K^4)$.

An example profile is presented for the Conestoga Creek Watershed near Lancaster, Pennsylvania (Figure 3). Here it is also shown how a profile can be defined mathematically by three parameters: the extent of overall information loss (A), the rate of information loss (B) and the asymptotic maximum conditional entropy (C). The parameters A and B reflect pattern in the spatial distribution, whereas C is directly related to the marginal (non-spatial) land cover distribution.

A landscape that is characterized by high dominance and patch coherence, leading to a very uneven distribution of land cover types both marginally and spatially, reveals the lowest entropy values for all resolutions when compared to other landscape types. Furthermore, the presence of very fine-grained categories causes the maximum attainable entropy to be depressed since these categories are 'washed out' as the resolution is degraded from a resampling filter. For example, starting with 8 land cover categories, the maximum conditional entropy is $\log(8^4) = 8.3$; whereas if a few degradations of resolution reduces the number of categories to 6, the maximum conditional entropy is reduced to $\log(6^4) = 7.2$.

As a dominant land cover, such as forest, becomes increasingly fragmented and replaced with other land cover categories, the marginal distribution becomes increasingly even. However, two landscapes can have the same marginal land cover distribution, yet have large differences in spatial distribution. This is expected to be captured by differences in the entropy profiles.

As the spatial distribution at the floor resolution becomes more evenly dispersed, with lower patch coherence and average patch size, the conditional entropy at the floor resolution increases. The profile also reaches a plateau rather quickly because the inherent pattern does not continue to change after just a few degradations of resolution; in which case the parameter A is expected to be small.

Meanwhile, if the spatial distribution at the floor resolution becomes more evenly distributed, but is characterized by high patch coherence and higher average patch size, conditional entropy at the floor resolution will be relatively small. However, the maximum attainable entropy will remain high since all, or most, land cover categories are maintained as the resolution is degraded.

Conventional single-resolution measurements

Besides obtaining a conditional entropy profile for each watershed, a suite of variables was also obtained using the FRAGSTATS software (McGarigal and Marks 1995). Also, certain land cover proportions were combined as follows. All forest cover types (conifer, mixed and deciduous) were summed to yield the proportion of 'Total Forest' cover, and annual herbaceous and perennial herbaceous land were summed to yield 'Total Herbaceous' cover. Terrestrial unvegetated land was also included to capture much of the urban land. The final set of single-resolution measurements are listed in Table 1.

Since all of the FRAGSTATS variables are explained in great detail in the corresponding manual (McGarigal and Marks 1995), as well as throughout the literature, such a discussion will be avoided here.

Multivariate clustering

Different combinations of the landscape variables were clustered by average euclidian distance after standardizing the variables. Manhattan distance made

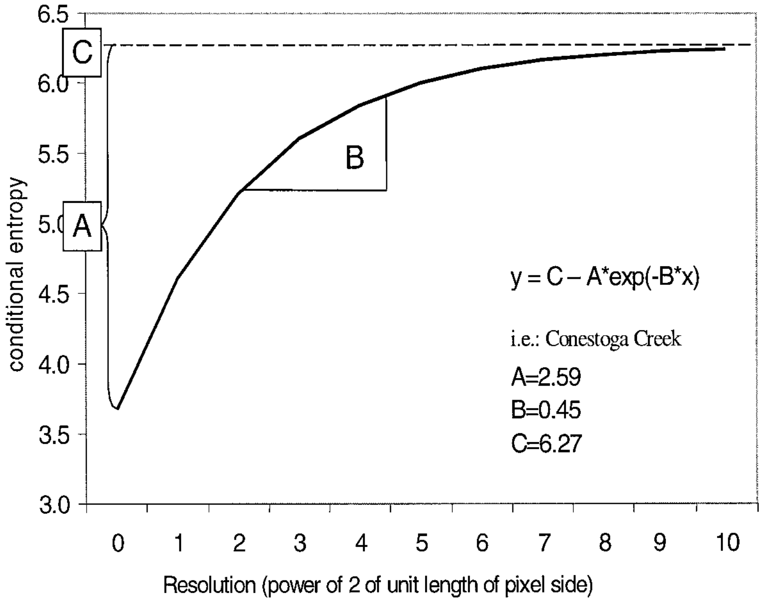

Figure 3. Example conditional entropy profile and associated parameters.

Table 1. Landscape variables measured for Pennsylvania watersheds.

Variable description	Code
Patch density	PD
Mean patch size	MPS
Patch size coefficient of variation	PSCV
Edge density	ED
Landscape shape index	LSI
Area-weighted mean shape index	AWMSI
Double-log fractal dimension	DLFD
Area-weighted mean patch fractal dimension	AWMPFD
Shannon evenness index	SHEI
Interspersion and juxtaposition index	IJI
Contagion[a]	CONTAG
Total forest cover	TOT.FOREST
Total herbaceous cover	TOT.HERB
Terrestrial unvegetated	TU

Diagonal pixels were included when determining patches.
[a]Pixel order preserved when measuring contagion.

Table 2. Different combinations of landscape variables used for clustering Pennsylvania watersheds.

Group	Description
1	All variables (11 single-resolution measurements, the profile variables A,B and C and 3 land cover proportions – total forest, total herbaceous and terrestrial unvegetated land)
2	All variables, excluding the profile variables A,B and C
3	Only A, B, C and the 3 land covers
4	Only the 3 land covers
5	Only the spatial pattern variables from group 1 (no land cover)

little difference once the variables were standardized, and the average linkage protocol was chosen over others for reasons of consistency and robustness. Also, model-based clustering was avoided because it assumes one has independent sample units and the watersheds of this study are a *population* of units that also exhibit high spatial dependence.

The different combinations of variables that were used for clustering the watersheds are listed in Table 2. The objectives of clustering were twofold. First, it was desired to determine which group of variables yields the most distinct watershed clusters with respect to apparent landscape fragmentation. In other words, for watersheds within a distinct cluster, as deciphered from a resulting dendrogram, do they appear to have similar land cover patterns that are in turn apparently different than watersheds in other clusters?

If the ‘best’ clustering results came from a large set of variables, then the second objective was to determine if a simpler subset could yield results that were not much different. At one extreme, it was desired to evaluate whether or not the simplest of measurements, Land cover proportions, alone could suffice for obtaining clear, sensible clusters with respect to landscape disturbance. Land cover proportions are generally expected to be very important; however, if measurements of spatial pattern substantially improved the cluster results, it was desired to determine if the conditional entropy profile variables alone were all that were necessary in addition to the land cover proportions.

Results and discussion

Relationships among the initial set of single-resolution spatial pattern variables from Table 1 are seen in Figure 4 for all of the 102 watersheds. The three land cover proportions and non-linear least squares regression estimates of the profile parameters A, B and C are plotted in Figure 5. An approximately uncorrelated subset of the pattern variables from Figure 4 was added to the plot in Figure 5 to visualize how the land cover proportions and profile variables are related to the larger set of pattern variables.

Pennsylvania presents a variety of different physiographic regimes that influence landscape-sculpting human activity in different ways. Therefore, analysis was done separately for each of three major physiographic provinces as discussed next, following Miller (1995).

The largest province is the *Appalachian Plateaus* that covers the western and northern part of Pennsylvania, encompassing about one half of the state's land area. This province is an upland area that is highly dissected by streams and rivers. Since the hilly topography discourages both agriculture and road building, human population densities are very uneven in the Appalachian Plateaus. In fact, most of the population lives within 100 miles of Pittsburgh where most of the bituminous coal is found. The northern edge of the Appalachian Plateaus has no coal and is largely isolated from major metropolitan centers, yet has the same rough topography that is not conducive to farming. Therefore, the northern edge is sparsely populated by people and contains some of the most pristine environmental conditions in the state (and greater region). These more pristine areas present a set of 'background' watersheds to which others can be compared.

The *Ridge and Valley* physiographic province encompasses the second largest area of Pennsylvania, curving from the southern border over to the eastern border. This province is very distinctive, from even a national perspective, due to its long, narrow valleys that are separated by long, narrow ridges. Since the ridges are steep and have thin soil, they remain mostly forested; however, the valleys are largely limestone, especially the Great Valley section along the southeastern edge of this province. The valleys are mostly cleared of forest to make way for farming and settlements.

The smallest of these three major provinces is the *Piedmont Plateau* in the southeast corner of Pennsylvania. This is a gently rolling, well-drained plain that is rarely more than 150 meters above sea level. Given that some of the best soil in the eastern United States occurs here, that the Piedmont encompasses Philadelphia, the nations fifth largest city according to the 2000 census, and lies in the New York City-to-Washington D.C. corridor, this province is highly developed with farms, urban centers and increasingly more suburban sprawl.

Clustering watersheds into common groups

The best clustering resulted from including all of the variables (group 1 in Table 2), regardless of some very high redundancy seen in Figures 4 and 5. The resulting dendrograms had very coherent structure that allowed ready separation of watersheds into distinct clusters. Watersheds of a common cluster appeared to have very similar land cover patterns that in turn appeared distinct from other clusters. This allowed labeling of the clusters with respect to relative degree and type of forest fragmentation. The resulting dendrograms appear in Figures 8 to **??**. Qualitative labeling of the clusters ranges from 'high', which relates to mostly forested watersheds, to 'very low', which relates to watersheds that are mostly agricultural with urban centers and extensive suburban development. These watersheds actually appear worse, with respect to forest cover, than those encompassing the cities of Philadelphia and Pittsburgh.

Given that five different groups of variables (Table 2) were evaluated for each of three physiographic provinces, one could elaborate for several pages on the clustering results. Instead, the key findings are summarized here.

– For the Piedmont Plateau, land cover proportions alone are sufficient to group the watersheds into clearly distinguishable landscape types. The only difference when clustering was done with the 3

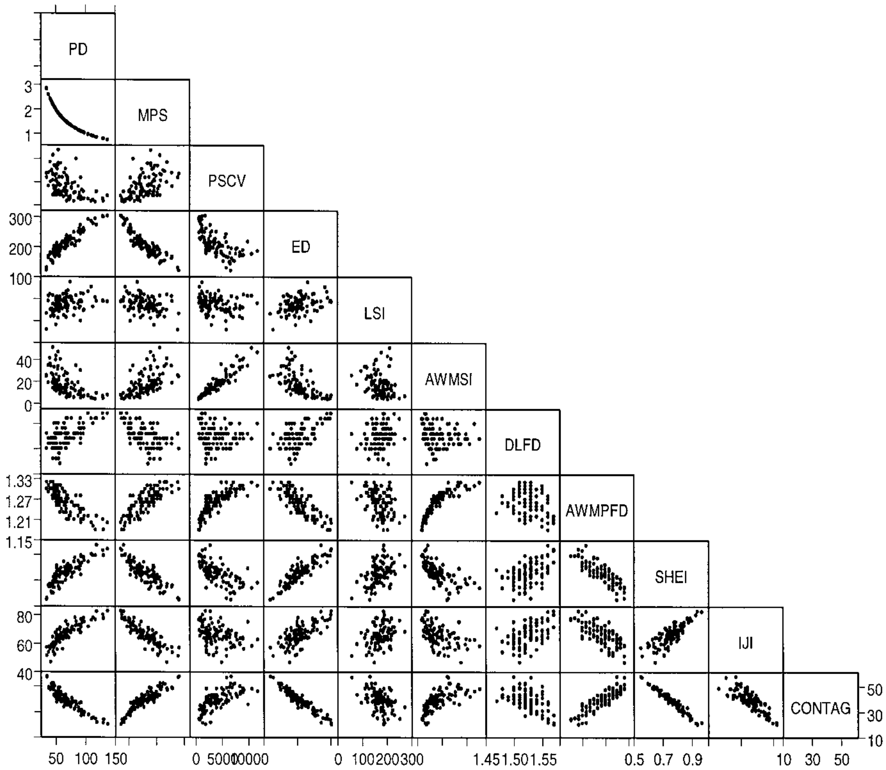

Figure 4. Pairwise scatterplots of all 102 watersheds for the landscape pattern variables in Table 1.

land cover proportions, compared to when all variables were used, is that White Clay Creek was drawn into the 'medium (suburban Philadelphia)' group, which is very reasonable. When all the pattern variables were included, but the land cover proportions were excluded, the 'medium' and 'very low' watersheds were mixed together, therefore reducing the ability to discriminate.

– For the Ridge and Valley, both land cover proportions and pattern variables were necessary to maintain reasonable clustering. When only the land cover proportions were used for clustering, two watersheds were distinctly separated off into their own group prior to splitting the remaining watersheds into clusters. These two watersheds (Conodoguinet Creek and Conococheague Creek) are very similar to the agricultural-dominated watersheds of the 'very low' group in the Piedmont Plateau that, as discussed above, were well distinguished by marginal land cover proportions alone.

– For the Ridge and Valley, when only the conditional entropy variables A, B and C were included with the land cover proportions, the resulting clusters were very similar to those obtained with the full set of variables. Some minor re-shuffling of a few watersheds occured, but nothing that yielded unreasonable results. However, when only the FRAGSTATS variables were included with the land cover proportions (no A,B and C values), watersheds in the clusters labeled 'upper and lower Appalachian Mountains' in Figure 7 became more randomly shuffled between these two clusters. This greatly reduced discriminatory ability because these original clusters represented geographically continuous physiographic sections within the Ridge and Valley Province that have internally consistent landscape patterns when visualized in the land cover maps. Therefore, the original clusters in Figure 7 had a rational interpretation, whereas this was lost when the conditional entropy variables were removed from the clustering. Consequently, one can argue that for the Ridge and Valley, just adding the conditional entropy values A, B and C to the land cover proportions may actually perform better than just adding the more conventional single-resolution variables.

– For the Appalachian Plateaus, the spatial pattern variables appear to have a stronger influence than

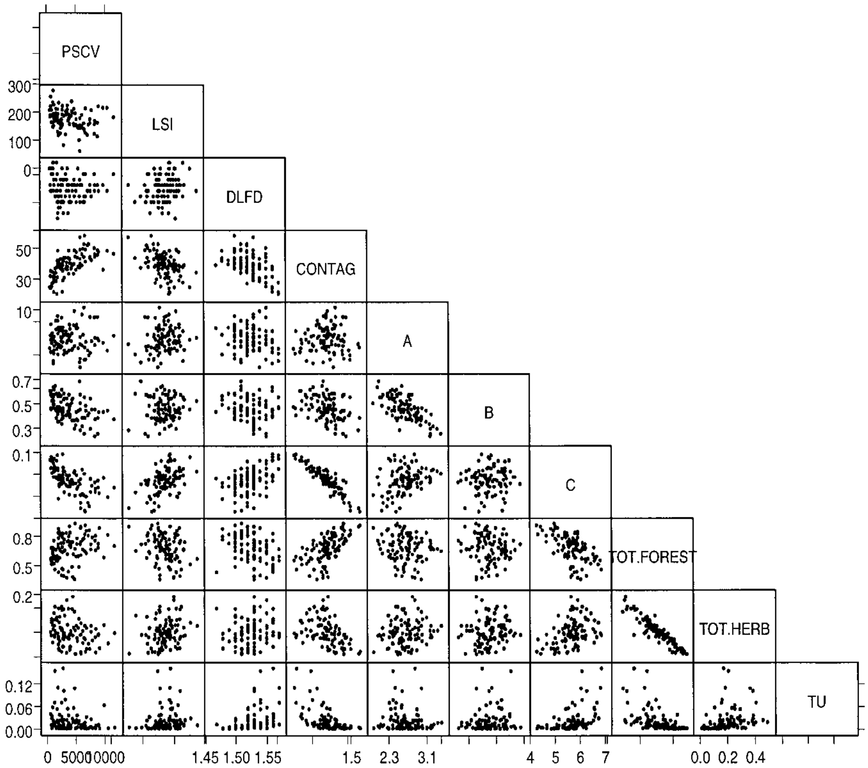

Figure 5. Pairwise scatterplots of all 102 watersheds for an approximately uncorrelated subset of pattern variables from Figure 4, along with the entropy profile variables A, B and C, and the land cover proportions listed as the last three entries in Table 1.

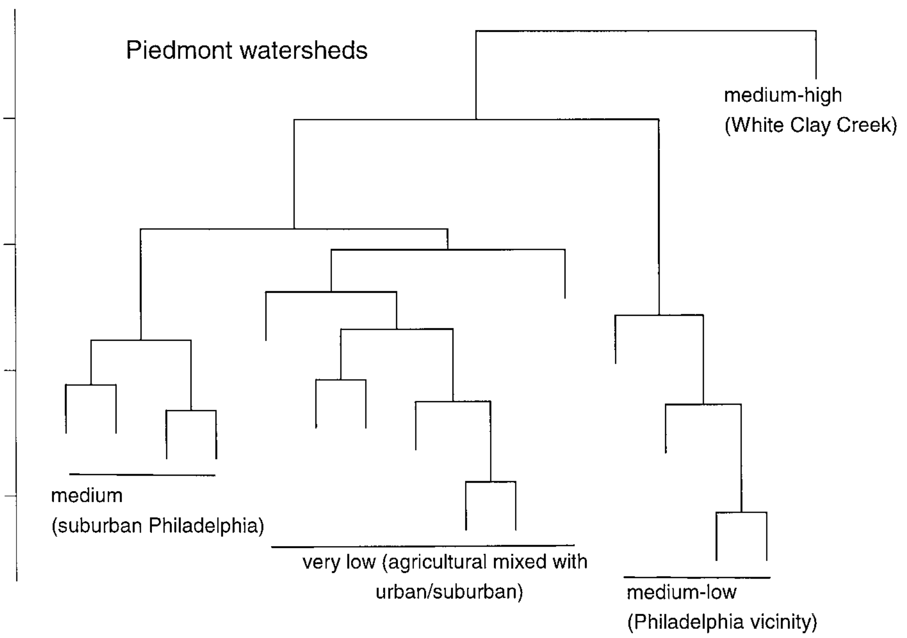

Figure 6. Cluster dendrogram of the Piedmont Plateau watersheds using the full set of landscape variables. Qualitative labeling of the clusters pertains to amount of forestation. Ordering of the watersheds, in reading from left to right, is simply an artifact of the software.

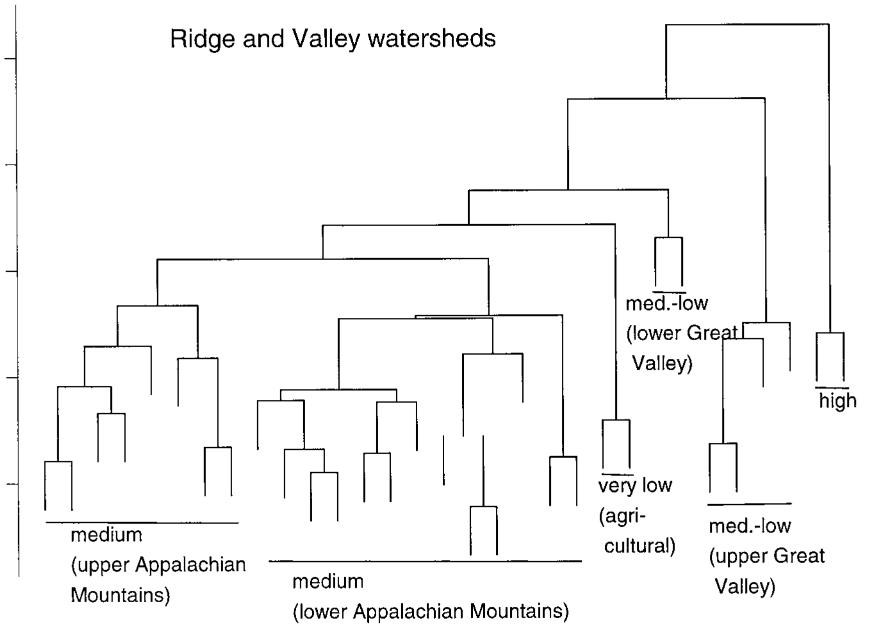

Figure 7. Cluster dendrogram of the Ridge and Valley watersheds using the same variables as for Figure 6. Note that 'Appalachian Mountains' and 'Great Valley' are physiographic sections within the Ridge and Valley physiographic province.

the marginal land cover proportions, precisely the opposite of findings from the Piedmont.

– For the Appalachian Plateaus, when only the conditional entropy variables A, B and C were included with the land cover proportions, one watershed (Turtle Creek) was distinctly separated into its own group that was very distant (in pattern space) from all other watersheds. This is reasonable because Turtle Creek is unique in that it encompasses the city of Pittsburgh, which is by far the largest urban center in the mostly rural Appalachian Plateaus. However, all of the remaining watersheds were basically split into 'medium-to-high' and 'medium-to-low' clusters that substantially sacrifices the greater detail obtained from clustering with all the original variables. When only the FRAGSTATS variables were included with the land cover proportions (no ABC), the resulting clusters were very similar to the original clusters obtained from using all the original variables. Therefore, one can argue that for the Appalachian Plateaus, the single-resolution variables out-perform the conditional entropy profile variables.
– One overall conclusion is that each physiographic province yielded its own unique properties.

Conditional entropy profiles

The conditional entropy profiles are reported in Figures 9 to 11 where they are coded with respect to the landscape clusters deciphered in Figures 6 to 8, respectively.

An initial, encouraging observation is that many profiles do separate from each other, thus indicating that they are responsive to changing landscape pattern in a way that can be readily graphed.

For the Appalachian Plateaus, a clear pattern is seen whereby the profiles rise to the top as the forest becomes increasingly fragmented. For both the Ridge and Valley and Piedmont Plateau, a pattern emerges whereby the profiles rise to the top, then 'collapse' to a more contagious state as the overall amount of non-forest land and the size of non-forest patches increases. This is entirely in line with earlier modeling results (Johnson *et al.* 1999). Apparently, the mostly fragmented watersheds of the Appalachian Plateaus are similar to more transitional (medium-to-low) watersheds of either the Piedmont or Ridge and Valley. It is interesting to note that the proportion of total forest cover for the highest overall profile in each of the three provinces is 0.50 for the Ridge and Valley (lower Lehigh River), 0.55 for the Piedmont (Wissahickon Creek) and 0.56 for the Appalachian Plateaus (Chartiers Creek). This indicates that there

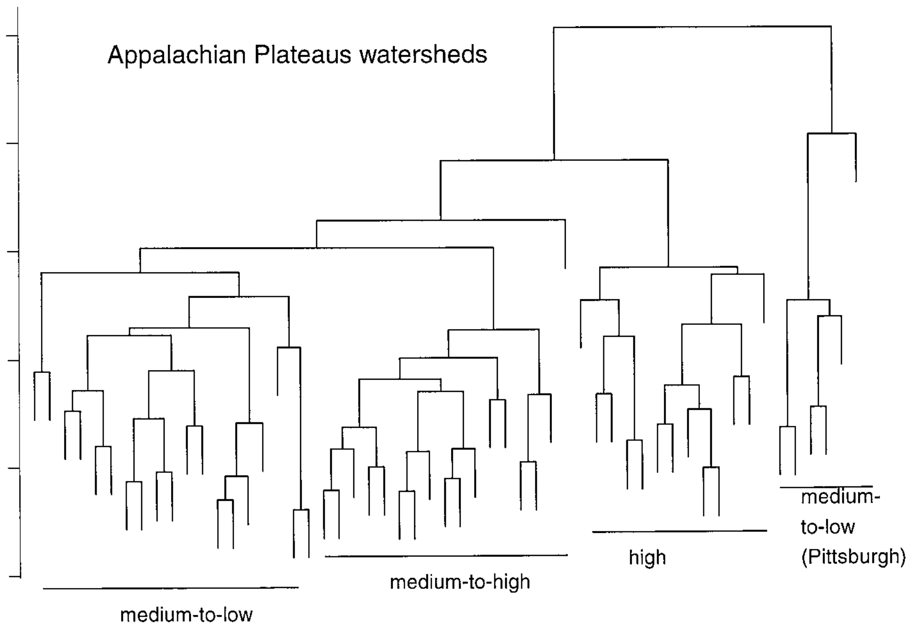

Figure 8. Cluster dendrogram of the Appalachian Plateaus watersheds using the same variables as for Figure 6.

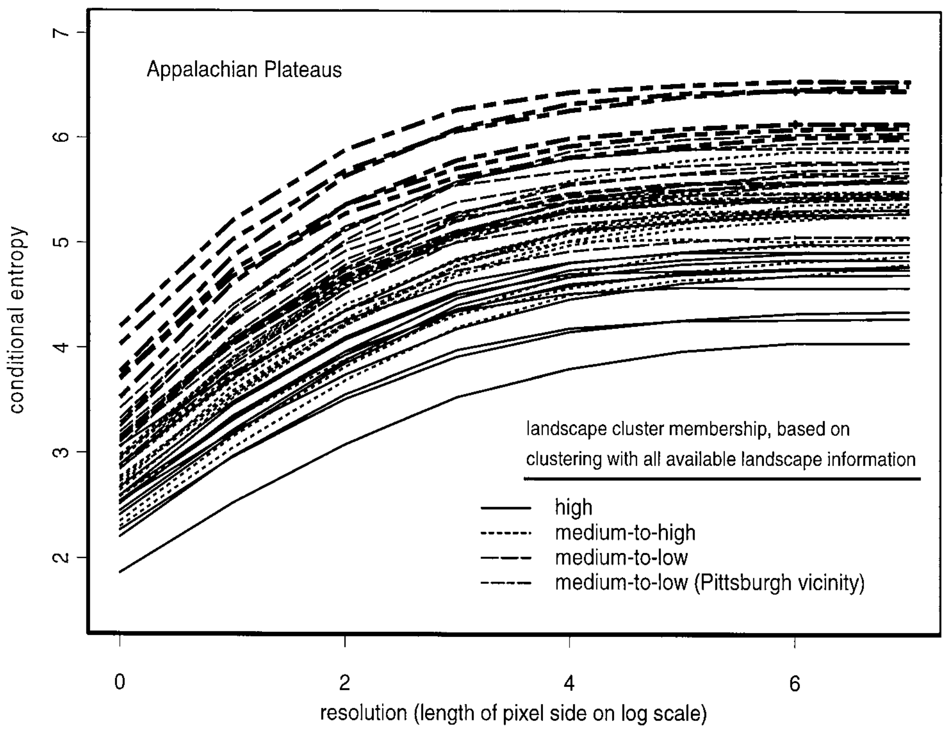

Figure 9. Conditional entropy profiles of watersheds in the Appalachian Plateaus physiographic province, coded according to membership in landscape cluster.

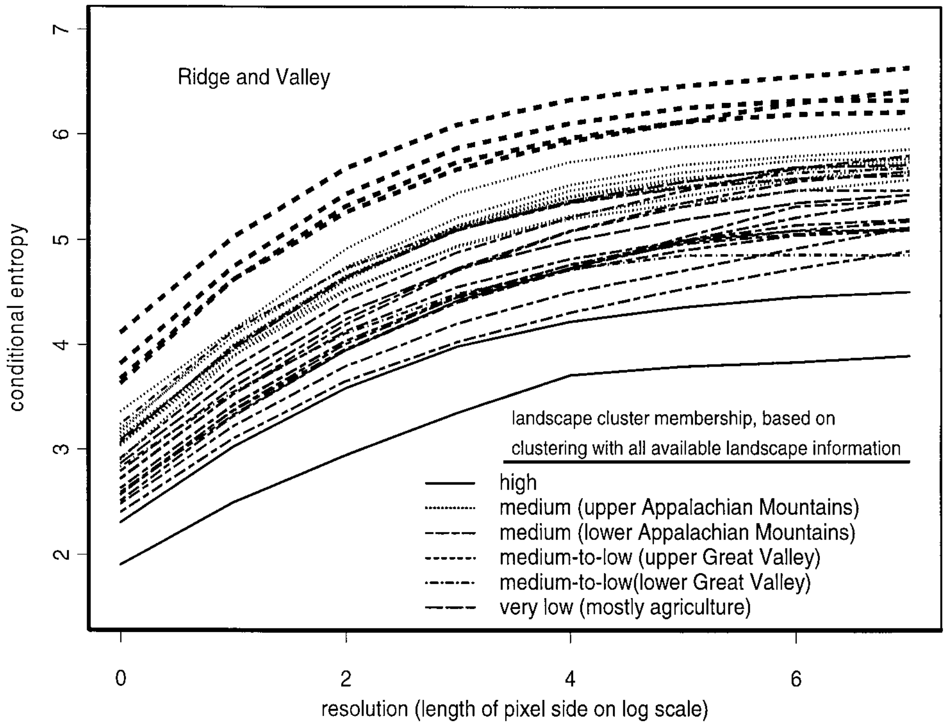

Figure 10. Conditional entropy profiles of watersheds in the Ridge and Valley physiographic province, coded according to membership in landscape cluster.

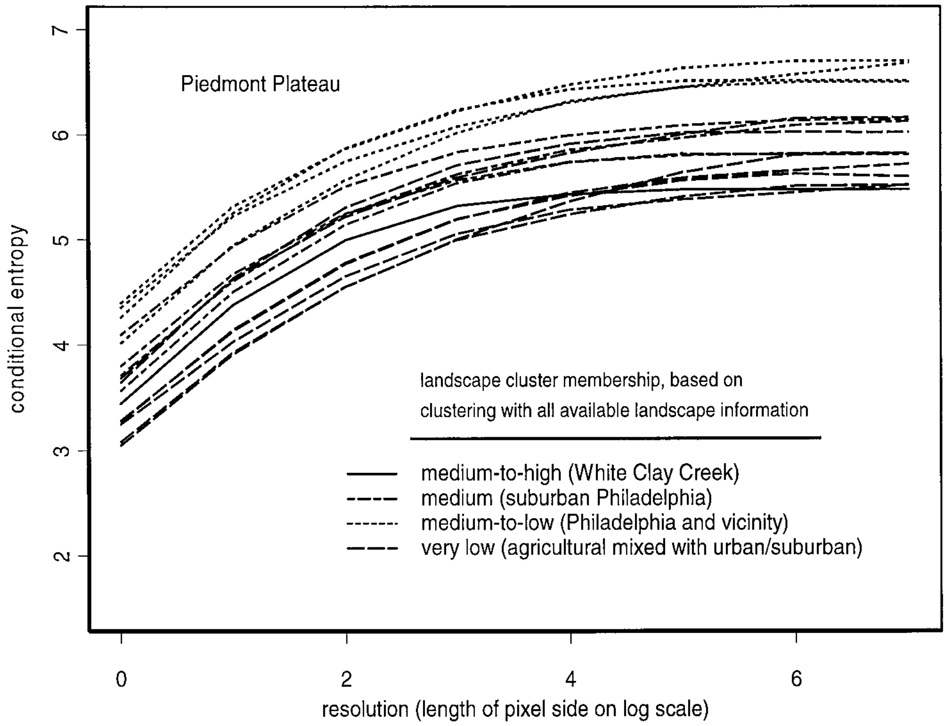

Figure 11. Conditional entropy profiles of watersheds in the Piedmont physiographic province, coded according to membership in landscape cluster.

is some consistency between the proportion of total forest cover and the maximum conditional entropy profile.

These profiles can also be used to compare fragmentation patterns over multiple resolutions between any two landscapes. If a profile from landscape **A** yields a higher conditional entropy than landscape **B** at every resolution, then we may say that landscape **A** is 'intrinsically more fragmented' than landscape **B** with respect to differing resolutions. On the other hand, if the two profiles cross at a particular resolution, then the two landscapes exhibit the same fragmentation pattern, as measured by conditional entropy, at that resolution. This is where additional information on land cover proportions would be needed to determine if these two landscapes are indeed different. If two landscapes have the same marginal (non-spatial) distribution of land cover, yet reveal conditional entropy profiles that are intrinsically separate, then it is indicated that we have two different spatial distributions of the land covers.

Conclusions

For Pennsylvania, USA, located in a temperate zone with a potential natural vegetation of predominantly hardwood (Mikan *et al.* 1994; Nowacki and Abrams 1992) or mixed conifer/hardwood (Whitney 1990) forest, the conditional entropy profiles provided a strong graphical tool with a sound quantitative basis for characterizing and monitoring landscapes. This achieves a goal that is similar to plotting different landscapes in 'pattern space,' as proposed by O'Neill, et al. (1996), which allows the ability to measure a landscape's 'distance' from some reference. However, conditional entropy profiles have the advantage of not only being multi-scalar, but also of being easily visualized since they are simple 2-dimensional plots.

Most variability among the landscapes is explained by the marginal land cover proportions and the asymptotic parameter C (as in Figure 3), which is in turn directly related to many single-resolution information-theoretic measurements (see Figures 4 and 5). Meanwhile, the other parameters describing a profile (A and B in Figure 3), appear to add new dimensions to a landscape characterization.

While these conditional entropy profiles proved useful for watersheds of the given scale within Pennsylvania, they remain to be evaluated on landscapes of much larger or smaller spatial extents, and also from other climatic and physiographic areas.

Acknowledgements

Prepared with partial support from the United States Environmental Protection Agency Cooperative Agreement Number CR-825506. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References

Benson, B.J and MacKenzie, M.D. 1995. Effects of sensor spatial resolution on landscape structure parameters. Landsc. Ecol. 10: 113–120.

Colwell, R.K. 1974. Predictability, constancy, and contingency of periodic phenomena. Ecol. 55: 1148–1153.

Costanza, R. and Maxwell, T. 1994. Resolution and predictability: An approach to the scaling problem. Landsc. Ecol. 9: 47–57.

Frohn, R.C. 1998. Remote Sensing for Landscape Ecology; New Metric Indicators for Monitoring, Modeling and Assessment of Ecosystems. Lewis, Boca Raton, FL, USA. 99 pp.

Johnson, G.D., Myers, W.L and Patil, G.P. 1999. Stochastic Generating Models for Simulating Hierarchically Structured Multicover Landscapes. Landsc. Ecol. 14: 413–421.

Johnson, G.D., Myers, W.L, Patil, G.P. and Taillie, C. 1998. Quantitative characterization of hierarchically scaled landscape patterns. 1998 American Statistical Association Proceedings of the Section on Statistics and the Environment, pp. 63–69.

Johnson, G.D., Myers, W.L, Patil, G.P. and Taillie, C. 1999. Multiresolution fragmentation profiles for assessing hierarchically structured landscape patterns. Ecol. Model. 116: 293–301.

Johnson, G.D. and Patil, G.P. 1998. Quantitative multi-resolution characterization of landscape patterns for assessing the status of ecosystem health in watershed management areas. Ecosys. Health 4: 177–187.

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: 123–131.

Jones, K.B, Riitters, K.H., Wickham, J.D., Tankersley, R.D. Jr., O'Neill, R.V., Chaloud, D.J., Smith, E.R. and Neale, A.C. 1997. An Ecological Assessment of the United States Mid-Atlantic Region. USEPA, ORD, EPA/600/R-97/130.

Kotliar, N.B. and Wiens 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59: 253–260.

McGarigal, K. and Marks, B. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA. 122 pp.

Mikan, C.J., Orwig, D.A. and Abrams, M.D. 1994. Age structure and successional dynamics of a pre-settlement-origin chestnut oak forest in the Pennsylvania Piedmont. Bull. Torrey Bot. Club 121: 13–23.

Miller, E.W. 1995. A Geography of Pennsylvania. The Pennsylvania State University Press. University Park, PA, USA 406 pp.

Myers, W. 1999. Remote Sensing and Quantitative Geogrids in PHASES [Pixel Hyperclusters As Segmented Environmental Signals], Release 3.4. Technical Report ER9710, Environmental Resources Research Institute, The Pennsylvania State University, University Park, PA, USA. 57 pp. and 2 diskettes.

Nowacki, G.J. and Abrams, M.D. 1992. Community, edaphic and historical analysis of mixed oak forests of the Ridge and Valley Province in central Pennsylvania. Can. J. Forest Res. 22: 790–800.

O'Neill, R.V., Gardner, R.H. and Turner, M.G. 1992. A hierarchical neutral model for landscape analysis. Landsc. Ecol. 7: 55–61.

O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Jackson, B.L., Jones, K.B. Riitters, K.H. and Wickham, J.D. 1996. Scale problems in reporting landscape pattern at the regional scale. Landsc. Ecol. 11: 169–180.

O'Neill, R.V., Johnson, A.R. and King, A.W. 1989. A hierarchical framework for the analysis of scale. Landsc. Ecol. 3: 193–205.

Patil, G.P. and Taillie, C.T. 1982. Diversity as a concept and its measurement. J. Am. Stat. Assoc. 77: 548–561.

Pielou, E. C. 1975. Ecological Diversity. Wiley-Interscience, New York, NY, USA.

Qi, Y. and Wu, J. 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. Landsc. Ecol. 11: 39–49.

Townshend, J.R.G. and Justice, C.O. 1988. Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. Int. J. Remote Sensing 9: 187–236.

Whitney, G.G. 1990. The history and status of the hemlock-hardwood forests of the Allegheny Plateau. J. Ecol., 78: 443–458.

Wiens, J.A. (1989). Spatial scaling in ecology. Funct. Ecol. 3: 385–397.