Landscape Ecology **12**: 199–212, 1997.

A multi-scale analysis of landscape statistics

Douglas H. Cain[1,2*], Kurt Riitters[2], and Kenneth Orvis[1]
[1]*Department of Geography, University of Tennessee, Knoxville, TN 37919, U.S.A.;* [2]*Tennessee Valley Authority, Norris, TN 37828, U.S.A.*

Keywords: Assessment, ecology, remote sensing, satellite

Abstract

It is now feasible to monitor some aspects of landscape ecological condition nationwide using remotely-sensed imagery and indicators of land cover pattern. Previous research showed redundancies among many reported pattern indicators and identified six unique dimensions of land cover pattern. This study tested the stability of those dimensions and representative pattern indicators across different types of land cover maps. The maps were derived from Landsat Thematic Mapper images of the Tennessee River and Chesapeake Bay watersheds, and they differed in resolution, number of attributes, and method of delineating landscape unit boundaries. A multivariate analysis of pattern metrics was conducted separately for each map, and the results were then compared among types of maps. Measures of land cover diversity, texture, and fractal dimension were more consistent than measures of average patch shape or compaction among the land cover maps.

Introduction

Concern over widespread loss and fragmentation of natural resources has led to calls for managing ecosystems in the context of their regional settings (*e.g.*, Holling 1978; Franklin 1993; Everett *et al.* 1994). It is a challenge to plan, manage, and assess changes at such scales, if only because biological systems are rarely studied with a regional, long-term perspective (Brown and Roughgarden 1990). Effective risk assessment and resource management requires a set of measurements or indicators to quantify biologically- and socially-relevant endpoints (Suter 1990). While there is an abundance of data for this purpose, particularly from remote sensing, regional biological systems are large and complicated, and there are few theories to suggest the critical measurements.

As part of a nation-wide effort to assess landscape ecological conditions, the Environmental Monitoring and Assessment Program–Landscapes (EMAP–L) of the Environmental Protection Agency is developing indicators which can be measured from remotely-sensed images – primarily land cover maps derived from satellite images (U.S. Environmental Protection Agency 1994). Many candidate indicators relate to the spatial patterning of land cover and various scaling relationships among landscape elements (Turner and Gardner 1991). Spatial pattern is a central feature of landscape ecology (Forman and Godron 1986); regional pattern often determines, and usually constrains finer-scale ecological condition (Turner 1989; Wiens 1989). For monitoring regional ecosystem conditions over time, EMAP-L must develop biologically-relevant indicators that are statistically independent, sensitive to real change but not to wild data, and estimable by remote sensing.

A large number of indicators can be calculated from mapped data (*e.g.*, Turner and Gardner 1991; Baker and Cai 1992; McGarigal and Marks 1994). But taken together, these indicators measured only six independent dimensions of pattern in a typical

**Present address:* The Sidwell Company, 28W240 North Ave., P.O. Box 920, West Chicago, IL 60186, U.S.A., (630) 231-0206.

collection of land cover maps (Riitters *et al.* 1995). More information is needed before a reduced set of indicators can be reliably implemented in the assessment effort. As a part of the overall EMAP–L project, the objective of this study was to examine statistical independence of pattern indicators for land cover maps of different scales and heritages, and to see whether the choice of landscape analyis unit ("watershed" versus "arbitrary rectangle") affects that independence.

Earlier research (Riitters *et al.* 1995) used U.S. Geological Survey Land Use Data Analysis (LUDA) maps which were created as vector coverages (Fegeas *et al.* 1983) and then rasterized at 200-meter resolution for analysis (Hunsaker *et al.* 1994). But EMAP–L intends to use raster satellite images (primarily TM/MSS) with 25 to 100 m resolution and fewer land cover attribute classes than the LUDA maps. Landscape elements (such as patches) are hand-drawn in vector maps but are computed for raster maps, and the scales of the two procedures may differ. A further complication is that the results of multivariate analyses are not necessarily independent of the choice of scale or stratification (Fotheringham and Wong 1991). Thus, a comparison of vector and raster maps was necessary.

Differences are expected for some indicators as the map parameters change; indeed, that is the basis for using indicators to assess real landscape changes. For example, indicators of fragmentation and dominance are sensitive to spatial resolution and the number of attribute classes (Turner 1990). The question addressed here is whether the *correlations* among indicators change for different types of maps. If the correlations do change, then separate sets of metrics will have to be chosen to summarize the pattern information in each type of map. Conversely, stability of the correlation structure will make it possible to use the same set of metrics for many different types of maps.

Test maps were created by altering Landsat Thematic Mapper (TM) land cover maps of the Chesapeake Bay Watershed (CBW) and the Tennessee River Watershed (TRW). The test maps were subdivided into analysis units, and 28 indicators were calculated for each unit. Factor analysis was used to summarize the correlation structure among the indicators, and to suggest subsets of indicators applicable for each type of map. The utility of these subsets for describing patterns in different types of maps was then evaluated by relative ability to discriminate among units, when the units had been grouped based on overall similarity of pattern.

Methods

Regional setting

Regional land cover patterns in the eastern United States reflect very broad-scale geophysical patterns, more or less modified by human influences. On regional vegetation maps, the visually obvious patterns are created by topography and human development. For example, ridge-and-valley zones are characterized by alternating corridors of ridge-top forest and bottomland non-forest vegetation, while mountainous zones contain rather large patches of forest, split by zones of human activity in river valleys. Humans dominate most low-elevation areas.

The Tennessee River Watershed drains over 10 million hectares in six states, westward from headwaters in the southern Appalachian Mountains (Fig. 1). There are four main physiographic regions – Appalachian Mountain, Ridge and Valley, Cumberland Plateau, and the Mississippi Embayment. The predominant land cover is forest (60%). Agricultural uses (crops and pasture) comprise 34%, with water (4.0%) and urban (1.4%) accounting for most of the rest.

The Chesapeake Bay Watershed drains almost 18 million hectares in six states, eastward from headwaters in the northern and central Appalachian Mountains (Fig. 2). The physiographic regions include the Appalachian Plateau, Ridge and Valley, Piedmont, and Coastal Plain. Forest accounts for 55% of the land cover in the watershed, followed by agriculture/pasture (33%), water (7.5%, mostly in Chesapeake Bay proper), and urban (4.5%) (U.S. Environmental Protection Agency 1994a).

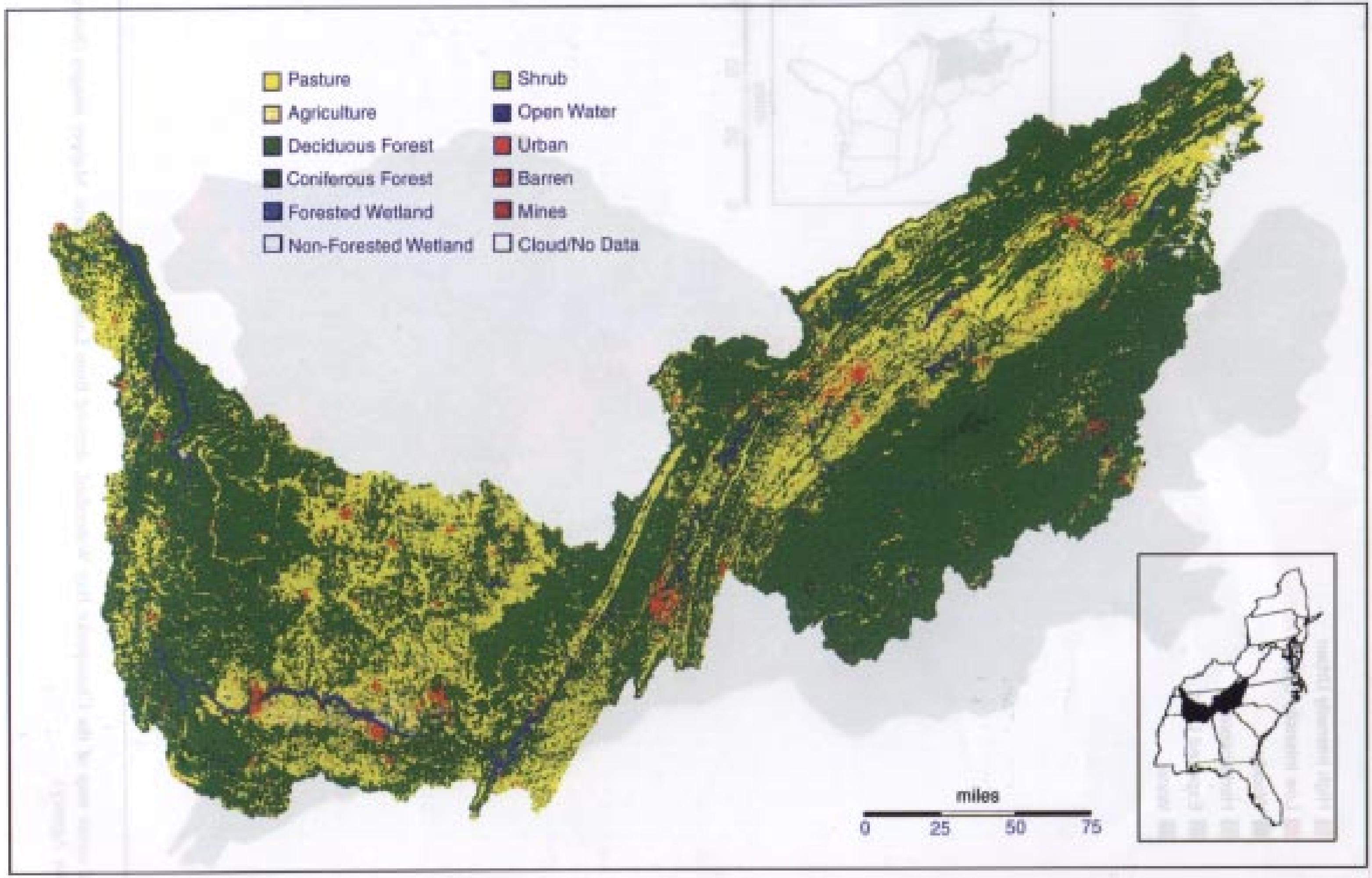

Fig. 1. 12-class land cover map of the Tennessee River Watershed, derived from Landsat Thematic Mapper images (Source: Tennessee Valley Authority).

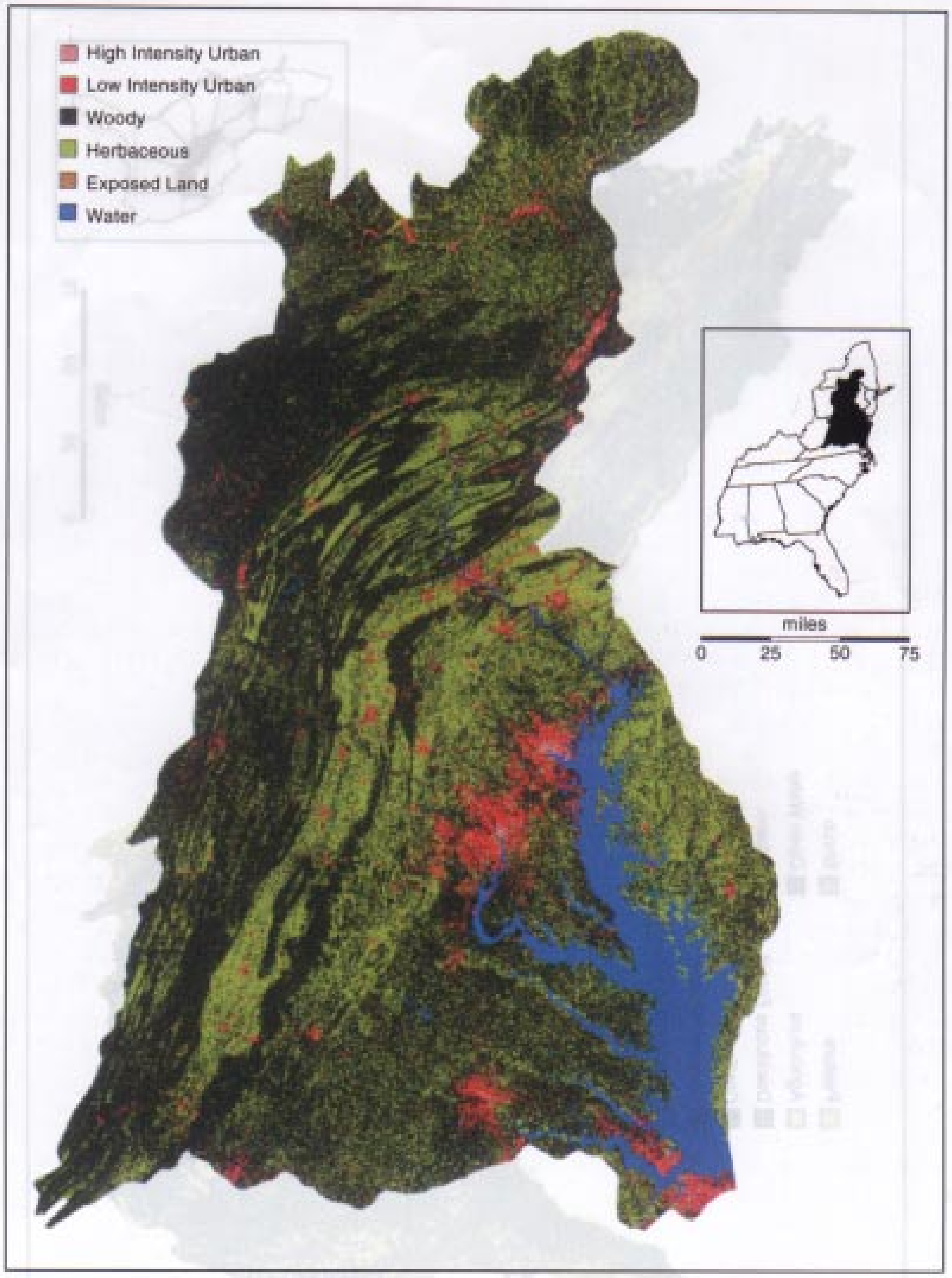

Fig. 2. 6-class land cover map of the Chesapeake Bay Watershed, derived from Landsat Thematic Mapper images (Source: Environmental Protection Agency).

Map generation

Land cover maps were available for both regions. Composite Landsat Thematic Mapper (TM) images had been classified by using both supervised and unsupervised classification procedures. The nominal resolution was 25 m for both maps. Twelve land cover classes were recognized in the TRW map, compared to six in the CBW map. Ground truth information came from aerial photography, maps, and other sources. The overall accuracy of land cover classification was reported as 85% for the TRW map and 80% for the CBW map. The high- and low-density urban classes in the CBW map were condensed into one urban class for this analysis.

Each map was split into "map sets" of rectangular equal-area analysis units. The units were ~1200 km^2 in the TRW map and ~1800 km^2 in the CBW map. 85 units were selected for analysis from the TRW map (denoted as T12_25), and 97 were selected from the CBW map (denoted as C5_25Q). The CBW map was also split into 46 sub-watersheds (C5_25W) defined by U.S. Geological Survey 8-digit hydrologic unit codes (HUC's, Seaber *et al.* 1987).

Test data sets were created as follows. The original TRW land cover map set (T12_25) was altered to create new sets with different spatial and attribute scales. To simulate a larger minimum mapping unit, a 5×5 majority rule filter was applied, and the 25-meter resolution was kept (T12_25M). To simulate a different grain size, the same filter was used but the output grain was set at 125 meters (T12_125M). A third map set was created by using the Arc/Info (ESRI 1992) nearest-neighbor algorithm which assigns the attribute class of the center cell to all cells in a 5×5 window (T12_125N). Finally, the attribute scale was varied by recoding the original 12-class map into just five classes (T5_25).

Land cover pattern indicators

Turner and Gardner (1991) provide an overview of pattern and scaling indicators of interest to landscape ecologists, including contagion, dominance, patch shape, and fractal dimension. Contagion quantifies the extent to which land covers coalesce to form larger patches, while dominance and diversity indicators measure the prevalence of one or a few land covers. Shape metrics quantify the shapes of landscape elements such as patches, sometimes in relation to standard shapes like circles or squares. Fractal models have been used to estimate land cover texture and the complexity of the perimeters of landscape elements. Additional information about pattern indicators may also be found in the literature of image processing (*e.g.*, Gonzalez and Woods 1992) or geography (*e.g.*, Lam and De Cola 1993).

Table 1 lists the indicators used in this study. The computing formulas and references for these indicators are given by Riitters *et al.* (1995). Briefly, most of the indicators are based on the frequencies of different attribute classes (*e.g.*, dominance), the frequencies of attribute class adjacencies (*e.g.*, contagion), or on the size and shape of contiguous clusters ("patches") of the same attribute class (*e.g.*, average patch size). A few others come from allometric relationships between two measures of landscape features (*e.g.*, fractal dimension). All 28 indicators were calculated for each of the analysis units.

Factor analysis

We conducted factor analyses on each of the data sets using the 28 indicators. Factor analysis (*e.g.*, Johnston 1980) is a multivariate procedure designed to reduce a large number of variables to a smaller set of "factors" which account for most of the variance among the original variables. Factors are typically extracted by applying principal components analysis to a standardized correlation matrix. A table of factor loadings shows which variables are grouped together on which common factors, and the degree of correlation between individual variables and the factors. The factors are interpreted as axes in state space, and the meanings of the axes are inferred from the variables which are most correlated with them. Highly-correlated variables are said to "load heavily" on that factor. Factors can be rotated in an attempt to account for additional variance, but non-orthogonal rotations produce correlated factors.

Table 1. Indicators of land cover pattern and structure used in this study.

NTYP: Number of Types
PMAX: Maximum Attribute Class Proportion
SIDI: Simpson Diversity of Attribute Classes
SIEV: Simpson Evenness of Attribute Classes
SHDI: Shannon Diversity of Attribute Classes
SHEV: Shannon Evenness of Attribute Classes
SUMD Sum of Diagonal Elements of Adjacency Matrix, *A*
SIDA: Simpson Diversity of Adjacency Matrix
SICO Simpson Contagion
SHDA: Shannon Diversity of Adjacency Matrix
SHCO: Shannon Contagion
FDDA: Area-Weighted Average of Fractal Dimension from Density-area Scaling
NPAT: Number of Patches
LPAT: Largest Patch
PSIZ: Average Patch Size
P005: Proportion of Area in which Patches are Greater than Five Pixels
PA-1: Average Perimeter-Area Ratio
PA-2: Average Adjusted Perimeter-Area Ratio
DSTA: Average Adjusted Perimeter-Area Ratio using Gardner's D-statistic
NASQ: Average Normalized Area, Square Model
RGYR: Average Radius of Gyration
PTRD: Average Patch Topology Ratio Dimension
ABRR: Average Bounding Rectangle Ratio
ABSR: Average Area Boxside Ratio
ACCR: Average Circumscribing Circle Ratio
ALAR: Average Area-by-Longest Axis Ratio
OCFC: Perimeter-Area Scaling-Pixels
OEFC: Perimeter-Area Scaling-Edges

Note: For explanation of indicators see Riitters *et al.* 1995. A Factor Analysis of Landscape Pattern and Structure Metrics. Landscape Ecology 10: 23–39.

Consider one type of map, for example. If all pattern indicators are mutually and highly correlated, then a single vector in pattern state space "explains" their variance, and a single factor will be identified. In that case, the choice of a particular indicator to represent the factor is easy because they all measure the same aspect of pattern. If there are two groups of indicators, then factor analysis will identify two vectors, and at least two indicators are needed to represent the state space, and so on. Ideally, representative indicators are strongly correlated with just one factor. In this way, factor analysis identifies both the number of different axes in pattern state space and the indicators which are associated with each axis.

The results of factor analyses for different test maps were compared by coefficients of congruence (Johnston 1980). The coefficient describes the similarity of the two factor patterns from different factor analyses by the following equation.

$$G_{ij} \frac{\sum_{k=1}^{n} L_{ki} L_{kj}}{\sqrt{\sum_{k=1}^{n} L^2_{ki} \sum_{k=1}^{n} L^2_{kj}}}$$

where

L_{ki} and L_{kj}

are the factor loadings for variable k in map types i and j, respectively;
n is the number of variables; and
G_{ij} is the coefficient of congruence between matrices i and j.

Large coefficients of congruence mean that the pattern indicators were extracted in similar fashions for two types of maps. The measure approaches a value of one when the loadings are proportional, and is effective in revealing similarities in different data sets (McDonald 1985). In a matrix of congruence statistics for two factor analyses, the rows and columns correspond to factors. High values on the diagonal indicate that the corresponding factors have similar loadings for each of the variables. Large off-diagonal values indicate situations where the same variables loaded on different factors in two types of maps. Similarly, the results of many factor analyses can be compared, one factor at a time, by constructing a second matrix which contains just the diagonal elements of the first matrix. The rows and columns of the new matrix are the different test data sets. If a given factor is consistently congruent across all test data sets, then all values of the new matrix will be large.

Table 2. Eigenvalues and cumulative proportion of variance explained by factor analysis of land cover pattern indicators.

Data set	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
T12_25						
Eigenvalue	14.296	5.635	3.474	1.342	1.135	.720
Cum. Variance	.511	.712	.836	.884	.924	.950
T12_125M						
Eigenvalue	13.948	4.890	3.533	1.465	1.221	1.028
Cum. Variance	.498	.673	.799	.851	.895	.932
T12_25M						
Eigenvalue	16.340	4.414	2.540	1.285	.866	.690
Cum. Variance	.605	.769	.863	.910	.942	.968
T12_125N						
Eigenvalue	13.581	4.822	3.484	1.952	1.273	.952
Cum. Variance	.485	.657	.782	.851	.897	.931
T5_25						
Eigenvalue	13.749	6.116	2.768	1.353	1.264	.931
Cum. Variance	.491	.710	.808	.857	.902	.935
C5_25Q						
Eigenvalue	11.522	8.040	2.889	1.645	1.016	.838
Cum. Variance	.412	.699	.802	.861	.897	.927
C5_25W						
Eigenvalue	13.144	7.444	2.336	1.522	1.018	.749
Cum. Variance	.469	.735	.819	.873	.909	.936

Note: See text for explanation of data sets

Selection of representative indicators

A representative set of indicators was selected as follows. For each type of map, groups of indicators were identified that loaded together on each of the extracted factors. The consistency of these groups for different types of maps was evaluated using congruence statistics. Individual metrics were selected to represent each group on the basis of a high correlation with a given factor and a low correlation with all others.

A discriminant analysis was done to test the adequacy of selected individual indicators in each map set. First, the factor scores were calculated for each anaysis unit. A factor score is an abstract quantity, simply a linear combination of indicator values, weighted by the corresponding correlation (loading) for that indicator with a particular factor. The analysis units were then grouped by using cluster analysis of the factor scores (using Ward's minimum variance algorithm, SAS 1982). Each group has similar values for all factor scores, and thus similar overall land cover patterns. Once clusters had been formed, the selected subset of individual indicators was tested for its ability to discriminate among the clusters. A stepwise method was used initially to test for statistical significance of the selected variables based on their discriminatory power as measured by Wilks' lambda statistic. The results are reported as mis-classification rates for the case where all selected variables were used in the discriminant function.

Results and discussion

From preliminary analyses[1], it was decided to extract six factors from each of the seven data sets using the principal components method applied to the standardized correlation matrix, followed by an orthogonal (varimax) rotation of axes. The six factors explained from 92% to 97% of the variation among indicators in the test data sets (Table 2). For all but one data set, the last one or two of

[1]Space does not permit showing the results of all seven factor analyses here. Details may be found in Cain, D.H. 1995. A multivariate analysis of metrics describing landscape pattern and structure. M.S. Thesis, Department of Geography, University of Tennessee, Knoxville, TN, 151 pp.

Table 3: Results of principal components analysis and varimax rotation for the Tennessee River Watershed original data set (T12_25).

Metric	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Communality
			Factor pattern after varimax rotation				
NTYP	-.160	.104	-.068	.092	.140	**.917**	.911
PMAX	**.889**	-.318	.054	.113	.113	-.144	.941
SIDI	**-.923**	.259	-.127	-.109	-.102	.153	.980
SIEV	**-9.31**	.246	-.120	-.125	-.120	.072	.977
SHDI	**-.912**	.212	-.245	-.044	-.077	.128	.960
SHEV	**-.917**	.169	-.220	-.092	-.131	-.131	.960
SUMD	**.940**	.149	-.028	.002	-.190	.057	.945
SIDA	**-.938**	.236	-.097	-.106	-.076	.142	.982
SICO	**.941**	-.232	.095	.110	.080	-.122	.982
SHDA	**-.948**	.153	-.212	-.034	-.031	.102	.980
SHCO	**.943**	-.105	.185	.083	.089	.171	.979
TMAS	**.936**	-.201	.076	-.120	-.039	-.019	.939
NPAT	**-.783**	**-.535**	-.079	-.019	.172	-.085	.942
LPAT	**.809**	-.236	.173	.147	.142	-.301	.872
PSIZ	**.799**	.474	.038	-.053	-.184	.025	.902
P005	**.686**	**.643**	.071	.057	-.146	.068	.919
PA-1	.347	**-.896**	.056	-.193	.057	.001	.967
DSTA	-.270	**.877**	.034	.340	-.153	.010	.983
PA-2	-.397	.356	-.328	**-.618**	.338	.221	.937
NASQ	.283	.152	.467	**.723**	-.302	.075	.941
RGYR	-.254	**.855**	-.208	-.173	-.087	.264	.947
PTRD	.394	-.106	**.874**	.012	-.009	-.094	.940
ABRR	.066	.432	.142	**.820**	-.186	.168	.946
ABSR	-.303	**.551**	**.662**	.191	-.228	-.148	.944
ACCR	.383	**-.615**	**.633**	.157	-.083	.013	.958
ALAR	.219	-.069	**.904**	.285	-.110	.013	.963
OCFC	.115	-.089	-.222	-.217	**.904**	.144	.955
OEFC	.205	**-.521**	.014	-.353	**.715**	.021	.951
						Sum	26.603
			Variance explained by each factor after rotation				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	
	12.892	5.118	3.156	2.207	1.918	1.311	

Notes: 1. See text for explanation of data sets
2. Numbers in bold indicate factor loadings greater than 0.5

the factors would be declared insignificant by the criterion that the eigenvalue be greater than 1.0 (*e.g.*, Johnson and Wichern 1982). All six factors were retained to facilitate comparisons among data sets.

The factor loadings are illustrated for the 12-class, 25-meter Tennessee River data set (T12_25) in Table 3. In this example, indicators of land cover dominance and contagion were highly correlated with the first factor but not the others. The second, third, and fourth factors are dominated by indicators of average patch shape (*e.g.*, perimeter-area ratios, compaction), while indicators of perimeter complexity and the number of attribute classes appear on the fifth and sixth factors, respectively. Note that the magnitude (but not the signs) of the loadings are important. Because most of the tested indicators are measures of texture, many of them were correlated with the first factor. In contrast, there was just one measure of the number of attribute classes, and consequently it alone was correlated with the last factor. This

Table 4. Summary of factor analysis results for seven data sets using principal components analysis with varimax rotation.

	Data set						
Metric	T12_25	T12_125M	T12_25M	T12_125N	T5_25	C5_25Q	C5_25W
NTYP	**6**	**5**	**5**	6	**6**	**6**	nhl
PMAX	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SIDI	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SIEV	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SHDI	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SHEV	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SUMD	**1**	**1**	**1**	**1**	**1**	4	**1**
SIDA	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SICO	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SHDA	**1**	**1**	**1**	**1**	**1**	**1**	**1**
SHCO	**1**	**1**	**1**	**1**	**1**	**1**	**1**
FDDA	**1**	**6**	**1**	**6**	**1**	1,**4**	1
NPAT	**1**	**1**	**1**	**1**	1	**2**	2
LPAT	**1**	**1**	**1**	**1**	**1**	4	1,6
PSIZ	**1**	**1**	**1**	**1**	1,5	**2**	**2**
P005	1,2	**1**	nhl	**1**	1,3	**2**	**2**
PA-1	**2**	3	**2**	**3**	**3**	**2**	**2**,3
DSTA	**2**	3,4	**4**	**3**	**3**	**2**	**2**,3
PA-2	4	**2**	**2**	**2**	1,4	**3**	**5**
NASQ	**4**	**2**	**2**	**2**	**2**	**3**	4
RGYR	**2**	**2**	**2**	**2**	**5**	**2**	**2**
PTRD	**3**	**3**	2,**4**	**4**	**2**	**2**	**3**
ABRR	**4**	**2**	**2**	**2**	2	3,4	**4**
ABSR	2,3	**3**	**2**	3,4	2,3	**2**	**3**
ACCR	2,3	**2**	**2**	**2**,4	**2**	**3**	2
ALAR	**3**	2,**3**	**2**	**4**	**2**	2	**3**
OCFC	**5**	**4**	**3**	**5**	**4**	2,5	**2**
OEFC	2,**5**	**4**	**3**	**5**	**4**	**2**	**2**

Notes: 1. See text for explanation of data sets
2. Bold numbers indicate a loading of 0.7 or higher
3. Small numbers indicate a loading between 0.5 and 0.7
4. nhl – no high loadings

emphasizes the role of factor analysis in partitioning orthogonal subsets of indicators; the particular order in which factors emerge is partly a result of the indicators which are included in the analysis.

Table 4 summarizes the factor loadings obtained for all seven data sets and identifies the factor(s) which each indicator was most highly correlated with in each data set. For example, the first column in Table 4 is comparable to the factor patterns for the T12_25 data set as shown in Table 3. Consistency of these factor-indicator associations across data sets suggested reasonable groupings of indicators. The first group (those primarily loading on the first factor) contains the texture (*e.g.*, contagion and diversity) indicators and (for the TRW data sets) the indicators related to patch size and density-area scaling. The latter indicators shift to other factors and are grouped with other indicators of patch shape and compaction in the CBW data sets.

The second group is comprised of indicators related to patch shape and compaction, composed primarily of indicators loading on factors 2–4. Included are the indicators PA-1, DSTA, PA-2, NASQ, RGYR, PTRD, ABRR, ABSR, ACCR, and ALAR. However, this group is not very stable, in the sense that individual indicators are correlated with different factors (usually the second or third factor) depending on the data set. It seems plausible that the second and third factors are measuring a similar dimension, and although it is

Table 5. Coefficients of Congruence for six factors among seven data sets.

Data set	T12_25	T12_125M	T12_25M	T12_125N	T5_25	C5_25Q	C5_25W
					Factor 1		
T12_25	1.000						
T12_125M	**-.911**	1.000					
T12_25M	**-.958**	**.872**	1.000				
T12_25M	**-.905**	**.999**	**.867**	1.000			
T5_25	**-.991**	**.913**	**.958**	**.908**	1.000		
C5_25Q	**-.885**	**.915**	**.862**	**.912**	**.913**	1.000	
C5_25W	**-.897**	**.926**	**.868**	**.924**	**.922**	**.993**	1.000
					Factor 2		
T12_25	1.000						
T12_125M	-.429	1.000					
T12_25M	-.376	**.818**	1.000				
T12_125N	-.334	**.991**	**.800**	1.000			
T5_25	.071	**.646**	**.799**	**.653**	1.000		
C5_25Q	**.804**	-.293	-.082	-.236	.393	1.000	
C5_25W	**.795**	-.377	-.200	-.318	.226	**.937**	1.000
					Factor 3		
T12_25	1.000						
T12_125M	**.655**	1.000					
T12_25M	-.015	-.242	1.000				
T12_125N	.003	**.632**	-.334	1.000			
T5_25	.127	**.608**	-.340	**.865**	1.000		
C5_25Q	**.607**	.361	-.136	-.083	-.110	1.000	
C5_25W	**.615**	**.939**	-.170	**.621**	**.641**	.358	1.000
					Factor 4		
T12_25	1.000						
T12_125M	-.455	1.000					
T12_25M	.325	-.477	1.000				
T12_125N	.220	-.298	**.627**	1.000			
T5_25	**-.724**	**.817**	-.243	-.206	1.000		
C5_25Q	-.086	-.016	-.148	-.229	-.248	1.000	
C5_25W	**.681**	-.311	.105	.269	-.412	**-.632**	1.000
					Factor 5		
T12_25	1.000						
T12_125M	-.069	1.000					
T12_25M	.213	**.686**	1.000				
T12_125N	**.920**	-.023	.189	1.000			
T5_25	-.319	.274	-.024	-.191	1.000		
C5_25Q	.423	-.033	.018	.427	-.280	1.000	
C5_25W	.292	.517	.354	.181	.302	.110	1.000
					Factor 6		
T12_25	1.000						
T12_125M	-.068	1.000					
T12_25M	.353	.172	1.000				
T12_125N	-.463	**.847**	.110	1.000			
T5_25	**.764**	-.093	-.214	-.453	1.000		
C5_25Q	**.808**	-.106	.009	-.478	**.694**	1.000	
C5_25W	.089	-.050	-.545	-.218	.344	.480	1.000

Note: Bold numbers indicate congruence of 0.6 or higher sometimes possible to discern two different parts of that dimension, the ability to draw that distinction depends on the data set. Some indicators display obvious similarities. For example, with one exception, PA-1 and DSTA move from the second factor in both original data sets to the third factor in the recoded or resampled data sets. The indicators PTRD, ABSR, and ALAR appear together for most of the data sets, but then their correlations with the factors were also somewhat low. PA-2, NASQ, and ABRR usually load on the same factors, strongly on factor 2 in all of the resampled data sets. These indicators may be robust to the different methods used to create test maps, but were less consistent when comparing test maps to the originals.

A third group is made up of the fractal estimators of perimeter complexity (OCFC, OEFC). They group together consistently, and were usually independent of other indicators (*i.e.*, they alone had high loadings on a given factor). When this group does contain unique information, it appears as the third, fourth, or fifth factor in importance.

The final group has just one member, the number of types or attribute classes (NTYP). This indicator consistently loaded highly, and by itself, on the fifth or sixth factor. Therefore, it is the least important for explaining the variance among indicators, but based on communality, it could be more important on the two low-resolution data sets and the TRW 5-class and CBW sub-watersheds data sets. "Communality" shows the proportion of variance that a variable has in common with other variables, and lower communality for these data sets indicates that NTYP is not as redundant a measure at low resolution or with only a few types of land cover.

Table 5 shows the diagonal elements of the congruence matrices for each factor in each data set. Factor 1 shows high congruence across all the map sets, indicating that variables are loading similarly on all of the first factors. Again note that the absolute value (not the sign) of the congruence statistic is important for these comparisons. The congruence matrix for factor 2 splits the original TRW data set and the two CBW data sets from another group made up of the resampled and recoded TRW data sets. This means that factor 2 contains different indicators for those two situations. There is little discernible pattern for the other four factors, which means that the specific factors that contain different indicators change across different test maps. In other words, even though the same indicators may be involved, the order of their importance changes for different maps. In summary, the congruence analysis confirms that the first factor is stable across all data sets, and strongly suggests that some similarities exist for the second factor for some of the sets. But most of the indicators are not stable with respect to their order of importance for various map comparisons.

The results of clustering the analysis units based on factor scores are illustrated for the map sets T12_25 and C5_25W (Fig. 3–4). Even though the cluster analysis did not consider spatial arrangement of the units, the clusters correspond roughly to physiographic regions. This suggests that overall pattern is associated with physiographic region. This interpretation is difficult, however, owing to the large grain size of the cluster maps relative to that of physiographic region maps (these are not shown here).

Selection of the indicators for discriminant analysis was somewhat arbitrary. To select a subset of indicators for the discriminant analysis, some indicators were rejected if they did not consistently have high loadings on most data sets. The indicators FDDA, P005, LPAT, PA-2, ABRR, ABSR, ACCR, and NPAT were rejected in this way. The number of attribute types (NTYP) was also rejected; although it almost always loaded strongly on the same factor, that factor explained only a small portion of the variance. In maps with many land cover types, NTYP could be much more significant. Measures of texture, compaction or shape, and fractal complexity emerged from the factor analyses with some consistency, so a subset of indicators was chosen to represent those axes. The texture measures PMAX and SHCO were selected because they were the two texture measures least correlated with each other. Based on the congruence analysis, the composite axes of shape and compaction was represented by DSTA and NASQ, which loaded most consistently and strongly on factors representing those dimensions. The fractal complexity axis was represented by OEFC.

These representative metrics correctly predicted

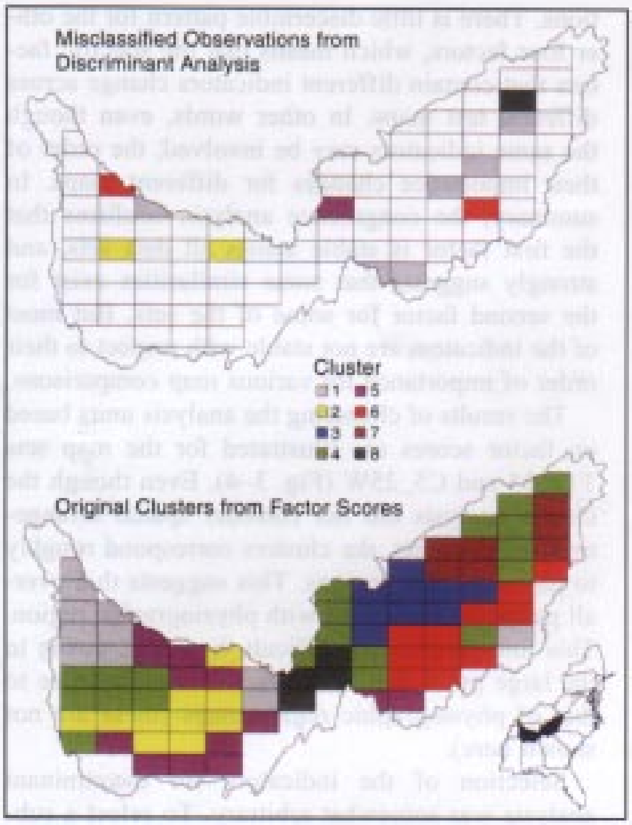

Fig. 3. Clusters and misclassified observations for the 12-class, 25-meter-resolution Tennessee River Watershed data set.

clusters created from the factor scores between 66% and 85% of the time, depending on the test data set. Figures 3 and 4 show the misclassified analysis units for the example data sets T12-25 and C5_25W. The best results were obtained for the hydrologic stratification of the CBW data set, and the worst were obtained for the rectangular splitting of the same data set. This suggests that land cover patterns are more homogeneous within sub-watersheds than within arbitrary rectangles. Together with the rough correspondence of clusters and physiographic regions, this is evidence that geophysical characteristics determine land cover pattern over large regions.

Conclusion

A statistical approach permits quantifying pattern, insofar as landscape ecologists know how to measure it. It is an open question whether such indicators are a complete basis for routine regional monitoring and assessment. No simple set of statistical pattern indicators can fully capture the biological complexity of ecosystems. But it seems reasonable, when developing relevant biological theory, to work with a set of indicators that is at least statistically independent. New theory will suggest new indicators to be tested, and may lend interpretation to the indicators in current use. The context lends biological relevance, but will not affect the correlations among indicators used to measure patterns in maps.

The research described here was in the context of broad-scale land cover assessment using satellite imagery. But the results could apply as well to other contexts, for example, fine-scale maps of soil parameters. At issue is not whether a given set of indicators describes all the important dimensions of the map, but rather, whether or not there is redundancy of a given set of indicators over a range of map types. Factor analysis is useful for making this set more efficient, but it is possible that additional orthogonal dimensions will be discovered if additional indicators are used.

An earlier study reported that six factors explained 87% of the variation among 26 pattern indicators for a set of LUDA maps (Riitters *et al.* 1995). The factors were interpreted as composite measures of average patch compaction, overall image texture, average patch shape, patch perimeter-area scaling, number of attribute classes, and large patch density-area scaling. Representative metrics selected from the factor analysis included average perimeter-area ratio (PA-1, contagion (SHCO), standardized patch shape (NASQ), patch perimeter-area scaling (a metric derived from OCFC), number of attributes (NTYP), and large patch density-area scaling (a indicator of "patch density", not used in this study).

The results of the present study are similar for the factors interpreted as texture, patch perimeter-area scaling, and number of attribute types, but patch shape and compaction did not always appear

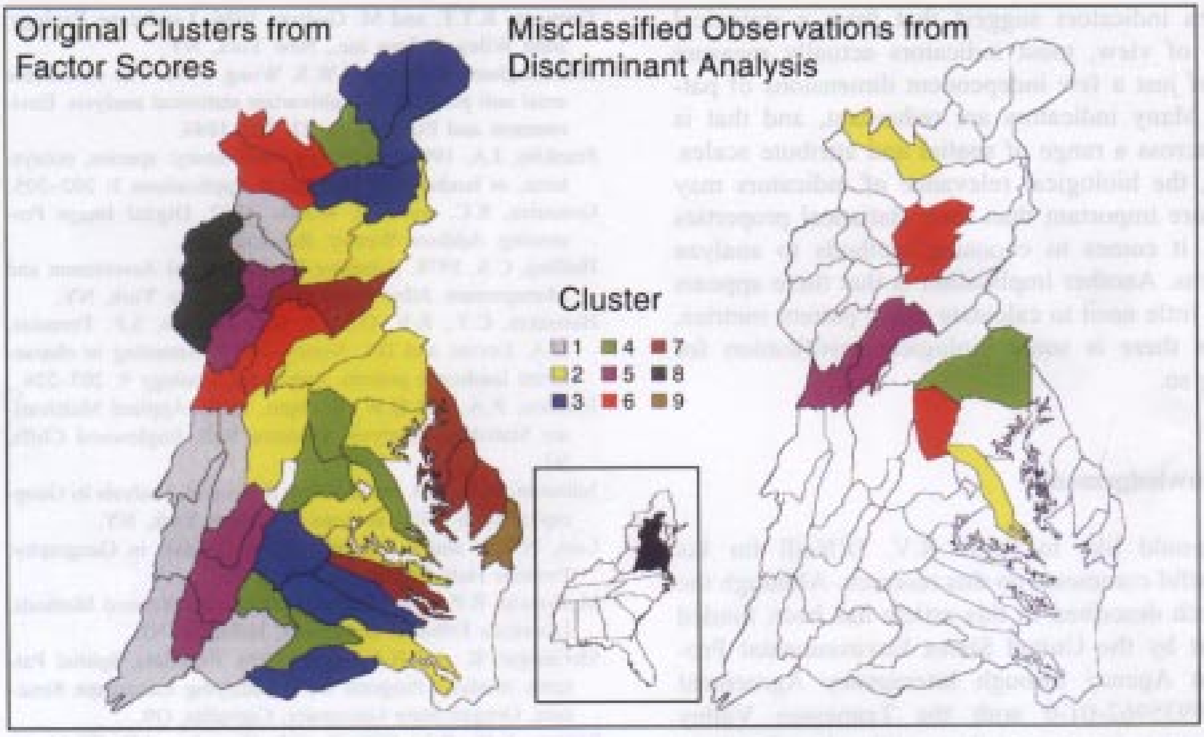

Fig. 4. Clusters and misclassified observations for the 5-class, 25-meter-resolution, sub-watershed-split Chesapeake Bay Watershed data set.

with patch shape and compaction sometimes load on just one factor, perhaps indicating that the same dimension is being measured in slightly different ways. A lack of consistency for these factors demonstrates that the correlations among some of the patch shape and compaction indicators may be unstable with respect to changes in the spatial or attribute scales of the input data.

In this study, four or five factors consistently appeared to be important. The most important factor was texture (measured by PMAX or SHCO) for all the data sets. This dimension seems to explain the majority of the variance among pattern indicators, regardless of spatial resolution, number of attributes, or stratification method. The second and third factors shared certain indicators which measured patch shape and compaction, but the indicators sometimes loaded on one factor, sometimes on the other, and sometimes on both. Another factor related to perimeter-area scaling was usually represented by the fourth or fifth factor sistently emerged alone as the sixth factor, but accounted for little variance.

Patch characteristics such as shape and compaction are presumably important aspects of a real landscape, even though their statistical correlations may be obscured by changing map scales. These indicators require further research, and in this case it may be better to evaluate them relative to their biological interpretations as opposed to statistical properties.

A subjective subset of five indicators (measuring texture, patch perimeter complexity, patch shape, and patch compaction dimensions) achieved classification accuracy rates of 66% to 85% when the goal of classification was to reproduce clusters obtained by considering factor scores representing the original, full set of 28 indicators. Incorporating an additional texture indicator (SHDA) improved the classification by up to 12% in some cases, but the improvement was inconsistent.

pattern indicators suggest that from a statistical point of view, most indicators actually measure one of just a few independent dimensions of pattern. Many indicators are redundant, and that is true across a range of spatial and attribute scales. Thus, the biological relevance of indicators may be more important than their statistical properties when it comes to choosing methods to analyze patterns. Another implication is that there appears to be little need to calculate many pattern metrics, unless there is some biological justification for doing so.

Acknowledgements

We would like to thank R.V. O'Neill for his insightful comments on this research. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through Interagency Agreement DW64935962-01-0 with the Tennessee Valley Authority, in cooperation with the Southern Appalachian Man and the Biosphere Program, it has not been subjected to Agency review and should not at this stage be construed to represent the views of the U.S. Environmental Protection Agency. Mention of trade names does not constitute endorsement or recommendation for use.

References

Baker, W.L. and Y. Cai. 1992. The r.le programs for multi-scale analysis of landscape structure using the GRASS geographical information system. Landscape Ecology 7: 291-302.

Brown, J.H. and J. Roughgarden. 1990. Ecology for a changing earth. Advances in Ecological Research.

ESRI. 1991. Arc/Info User's Guide. Environmental Systems Research Institute, Inc., Redlands, CA.

Everett, R., P. Hessburg, M. Jensen and B. Bormann. 1994. Eastside forest ecosystem health assessment. Volume I: Executive summary. General Technical Report PNW-GTR-317. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.

Fegaeas, R.G., R.W. Claire, S.C. Guptil, K.E. Anderson and C.A. Hallam. 1983. Land use and land cover digital data. Geological Survey Circular 895-E. U.S. Geological Survey, Washington, DC.

Formann, R.T.T. and M. Godron. 1986. Landscape Ecology. John Wiley & Sons Inc., New York, NY.

Fotheringham, A.S. and D.W.S. Wong. 1991. The modifiable areal unit problem in multivariate statistical analysis. Environment and Planning A. 23: 102–1044.

Franklin, J.A. 1993. Preserving biodiversity: species, ecosystems, or landscapes? Ecological Applications 3: 202–205.

Gonzalez, R.C. and R.E. Woods. 1992. Digital Image Processing. Addison-Wesley, Reading, MA.

Holling, C.S. 1978. Adaptive Environmental Assessment and Management. John Wiley and Sons, New York, NY.

Hunsaker, C.T., R.V. O'Neill, B.L. Jackson, S.P. Timmins, D.A. Levine and D.J. Norton. 1994. Sampling to characterize landscape pattern. Landscape Ecology 9: 207–226.

Johnson, R.A. and D.W. Wichern. 1982. Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliffs, NJ.

Johnston, R.J. 1980. Multivariate Statistical Analysis in Geography. John Wiley & Sons Inc., New York, NY.

Lam, N.S.N. and L. DeCola. 1993. Fractals in Geography. Prentice Hall, Englewood Cliffs, NJ.

McDonald, R.P. 1985. Factor Analysis and Related Methods. Lawrence Erlbaum Associates, Hillsdale, NJ.

McGarigal, K. and B.J. Marks. 1994. Fragstats, Spatial Patterns Analysis Program for Quantifying Landscape Structure. Oregon State University, Corvallis, OR.

Riitters, K.H., R.V. O'Neill, C.T. Hunsaker, J.D. Wickham, D.H. Yankee, S.P. Timmins, K.B. Jones and B.L. Jackson. 1995. A Factor Analysis of Landscape Pattern and Structure Metrics. Landscape Ecology 10: 23–39.

SAS Institute. 1982. User's Guide: Statistics. SAS Institute, Cary, NC.

Seaber, P.R., F.P. Kapinos and G.L. Knapp. 1987. United States Geological Survey Water Supply Paper 2294. U.S.G.S., Denver, CO.

Suter, II, G.W. 1990. Endpoints for regional ecological risk assessments. Environmental Management 14: 19–23.

Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology and Systematics 20: 171–197.

Turner, M.G. 1990. Spatial and temporal analysis of landscape patterns. Landscape Ecology 4: 21–30.

Turner, M.G. and R.H. Gardner. 1991. Quantitative Methods in Landscape Ecology: The Analysis and Interpretation of Landscape Heterogeneity. Springer-Verlag, New York, NY.

United States Environmental Protection Agency. 1994. Landscape Monitoring and Assessment Research Plan. U.S. EPA 620/R-94-009. Office of Research and Development, Washington, DC 20460.

United States Environmental Protection Agency. 1994a. Chesapeake Bay Watershed Pilot Project. U.S. EPA 620/R-94/020. Office of Research and Development, Washington, DC 20460.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385–397.