J Geograph Syst (2002) 4:69–83

Journal of Geographical Systems

Area-based tests for association between spatial patterns

Susan L. Maruca, Geoffrey M. Jacquez

BioMedware, Inc., 516 N. State St., Ann Arbor, MI 48104, USA
(e-mail: maruca@biomedware.com)

Received 1 June 2001 / Accepted 25 October 2001

Abstract. Edge effects pervade natural systems, and the processes that determine spatial heterogeneity (e.g. physical, geochemical, biological, ecological factors) occur on diverse spatial scales. Hence, tests for association between spatial patterns should be unbiased by edge effects and be based on null spatial models that incorporate the spatial heterogeneity characteristic of real-world systems. This paper develops probabilistic pattern association tests that are appropriate when edge effects are present, polygon size is heterogeneous, and the number of polygons varies from one classification to another. The tests are based on the amount of overlap between polygons in each of two partitions. Unweighted and area-weighted versions of the statistics are developed and verified using scenarios representing both polygon overlap and avoidance at different spatial scales and for different distributions of polygon sizes. These statistics were applied to Soda Butte Creek, Wyoming, to determine whether stream microhabitats, such as riffles, pools and glides, can be identified remotely using high spatial resolution hyperspectral imagery. These new "spatially explicit" techniques provide information and insights that cannot be obtained from the spectral information alone.

Key words: Hyperspectral imagery, spatial statistics, pattern recognition, stream habitat

JEL classification: C15, C69, C89

This manuscript benefited substantially from the critical reading of 3 anonymous referees. We wish to thank Andrew Marcus for his thoughtful criticisms and comments. This research was funded by a grant from the National Institute of Environmental Health Sciences to G. M. Jacquez and BioMedware. The views expressed in this publication are those of the authors, and do not necessarily reflect the official views of the NIEHS. We thank Bob Crabtree and the Yellowstone Ecosystem Studies group for providing the test data sets.

1 Introduction

A recent workshop on high spatial resolution hyperspectral (HSRH) imagery in exposure assessment (see Jacquez et al. in this issue) identified the development and application of "spatially explicit" methods as a critical research need. Spatially explicit methods make use of both spatial and spectral information. Consider the spatially-referenced data model $\{x_i, y_i, a_{i1},\ldots,a_{ik}\}$. Here x_i, y_i is the coordinate of location (e.g. pixel) *i*, and $a_{i1},\ldots,a_{ik}$ are observations on *k* variables (e.g. spectral bands) at that location. While purely spectral methods explore multivariate covariance in the spectral dimension ($a_{i1},\ldots,a_{ik}$), spatially explicit techniques quantify how this covariance changes through geographic space. Because spatial relationships mediate most if not all physical, geochemical, biological and ecological processes, it seems reasonable to suggest that spatially explicit techniques might reveal relationships that are not apparent in the spectral dimension alone.

While spectral analysis methods for HSRH imagery are numerous and well-developed (see for example Aspinall et al. in this issue), spatially explicit methods for HSRH image analysis are relatively few. Although spatial analysis has a rich and robust repertoire of techniques (e.g. in the traditions of spatial autocorrelation analysis, autoregressive modeling, geostatistics and so on), these methods were not developed for HSRH imagery. The extension of these techniques to hyperspectral image analysis is considered by others in this special issue (see for example the works by Griffith, Goovaerts and Rogerson). In this paper we consider the analysis of HSRH imagery as a problem in pattern recognition and comparison. The first step is to identify patterns on an image; the second step is to relate the observed patterns to patterns in other data sets. Scientific questions that may be considered in this framework include "are landscape features associated with bedrock geology?" "Are vegetation patches associated with edaphic characteristics", "Are land-use patterns associated with forest types" and so on.

We build on the field-object paradigm that is now commonly used as a foundation for data structures and algorithms in geocomputation (Egenhofer et al. 1999, Mark et al. 1999). We expect that field-object transformations will prove useful for data reduction on HSRH imagery, by identifying relevant collections of "signals" (i.e. objects) on these large, multivariate raster images (i.e. fields). The meaningful association of extracted features (such as patches) to real-world phenomena has direct implications for our ability to analyze and extract geospatial knowledge from HSRH imagery. In this paper we develop statistical tests for assessing the correspondence between features extracted from HSRH images and other spatial objects. Existing approaches to this problem include boundary overlap statistics (Jacquez 1995, Fortin et al. 1996) that determine whether boundaries for two or more variables coincide, or overlap, to a significant extent (for applications see Jacquez 1995, Fortin et al. 1996, Fortin 1997, Hall and Maruca 2001). These statistics were designed to be used with difference boundaries, which are geographic zones of rapid spatial change in the underlying variable(s) (Jacquez et al. 2000). Difference boundaries are distinguished from area (polygonal) boundaries, in that (1) difference boundaries indicate local changes while area boundaries reflect larger-scale spatial pattern, and (2) difference boundaries do not enclose or represent an area. The primary

limitation to these difference boundary distance-based statistics is that, in relying on boundaries as local, discrete objects, they are not useful for analyzing the overlap of one set of polygonal boundaries with another, especially when each of the two sets represents an exhaustive partition of the geographic study space. How would one measure the 'distance' between two polygons that intersect? Further, in many cases we are more interested in determining whether the *areas themselves* overlap, and not necessarily whether the enclosing polygon lines overlap. An area-based approach would be more effective at quantifying the 'match' between two sets of polygons.

There is a substantial body of prior work on the modeling of spatial and spatio-temporal relationships among polygons (for access to this literature see Bonham-Carter 1994 and the introduction to Sadahiro and Umemura, 2000). Sadahiro and colleagues (Sadahiro 2000, Sadahiro and Umermura 2000) developed two techniques; one for assessing polygon overlap within an exploratory change detection context, and another for the space-time analysis of polygon distributions. To our knowledge, none of the existing techniques are developed within a probabilistic framework, and thus do not lend themselves to inferential statistics for assessing association between two spatial patterns. Here we are concerned with the development of techniques for the probabilistic assessment of polygon overlap. This approach supports hypothesis testing and statistical inference and evaluates whether an observed amount of overlap is "unusual" relative to a null or neutral spatial model representing a biologically reasonable null hypothesis.

Consider the two sets of polygons in Fig. 1A–B. One set of polygons is regular, and the other set is irregular. Our intention was to develop a method that distinguishes instance (A), where the two sets of polygon areas coincide, from instance (B), where they do not.

2 Methods

2.1 Area overlap statistics

We begin with two sets of polygons, **I** and **J**, each comprised of N_I and N_J polygons and obtained as an exhaustive partitioning of the same

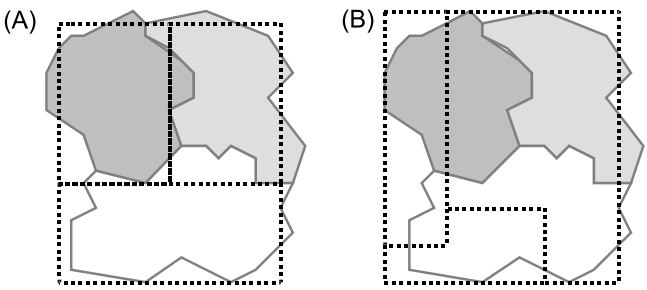

Fig. 1. Examples of hypothetical, simplified partitions that overlap (**A**) and don't overlap (**B**). Each pair of partitions (**A** and **B**) contains one set of irregularly shaped polygons (*solid outline*) and one set of rectangular polygons (*dashed outline*). Area overlap analysis uses appropriate null spatial models to distinguish (**A**) from (**B**)

geographic space. The idea is to see how well the two sets overlap each other with respect to the areas delineated by the polygons. For polygon i in set **I** and polygon j in set **J**, we calculate the following quantity, called the *relative area overlap*:

$$a_{ij} = \frac{a_{(i \cap j)}}{a_{(i \cup j)}} \tag{1}$$

where $a_{(i \cap j)}$ is the area of intersection and $a_{(i \cup j)}$ is the area of union for polygons i and j. Relative area overlap for non-intersecting polygons is zero, and increasing values represent better overlap, with a maximum value of 1 for perfectly overlapping polygons (where $a_{(i \cap j)} = a_{(i \cup j)}$). For example, the relative area overlap for the polygon pair in Fig. 2A is 0.05, while a_{ij} for the pair in Fig. 2B is 0.52.

For each polygon i in **I** we can then find the polygon in **J** that i overlaps best with, by finding the maximum value of a_{ij} over all polygons in **J** (called the *maximum relative area overlap*):

$$A_i = \max(a_{i\bullet}) \tag{2}$$

We then define an area overlap statistic for set **I** as the average maximum relative area overlap:

$$A_I = \frac{\sum_{i=1}^{N_I} \max(A_{i\bullet})}{N_I} \tag{3}$$

Similarly, for set **J**:

$$A_J = \frac{\sum_{j=1}^{N_J} \max(A_{\bullet j})}{N_J} \tag{4}$$

Further, we can calculate a simultaneous area overlap statistic as:

$$A_{IJ} = \frac{\sum_{i=1}^{N_I} \max(A_{i\bullet}) + \sum_{j=1}^{N_J} \max(A_{\bullet j})}{N_I + N_J} \tag{5}$$

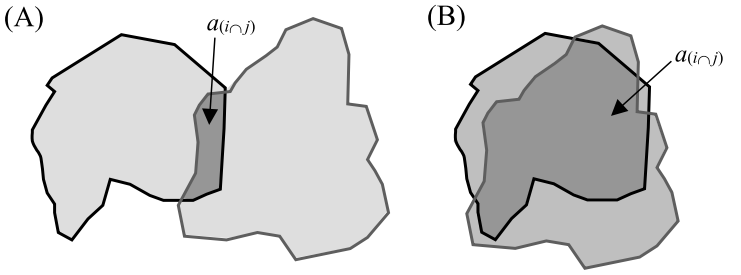

Fig. 2. Polygon pair arrangements showing (**A**) low area overlap and (**B**) high area overlap. For any polygon pair, relative area overlap is defined as the ratio of the area of intersection, $a_{(i \cap j)}$ (*darker areas*), to the area of union, $a_{(i \cup j)}$ (*all shaded areas*)

The statistic A_I is a measure of how well the polygons in **I** overlap, or 'match up with', the polygons in **J**. Likewise, A_J is a measure of how well the polygons in **J** overlap with those in **I**. A_{IJ} is a general (or bi-directional) measure of overlap between the two sets of polygons.

We expect these statistics to be most useful for cases where the two partitions contain roughly the same number of polygons (that is, the partitioning was conducted on similar spatial scales), and the variance in polygon size is relatively low. However, in the real world, the researcher may not have control over the scale of the partitioning, which could result in two partitions with very different numbers of polygons. Further, a given partition may very reasonably contain polygons of drastically different sizes. In particular, edge effects may be present so that polygon size may decrease near the margins of landscape boundaries and other features. In these cases, we propose calculating the average maximum relative overlap (e.g. A_I) as a weighted average, where the weighting factor is the area of the focus polygon (a_i for polygon i in set **I**; a_j for polygon j in set **J**). Focus polygons are selected in one of two frameworks, a *global* analysis and a *local* analysis. When calculating a global overlap statistic one considers each polygon in the segmentation (data layer) by turn. Note that the order in which these are considered has no impact on the value of the test statistic. One also can evaluate the overlap statistic within specific portions of the study area. This is a local overlap statistic in which the focus polygons are those that comprise the specific portion of the study area under consideration. Again, the order in which this subset of local polygons is considered makes no difference. In this scenario, the statistics would be calculated as follows:

$$A_I = \frac{\sum_{i=1}^{N_I} [a_i \max(A_{i\bullet})]}{\sum_{i=1}^{N_I} a_i} \tag{6}$$

$$A_J = \frac{\sum_{j=1}^{N_J} \left[a_j \max\left(A_{\bullet j}\right)\right]}{\sum_{j=1}^{N_J} a_j} \tag{7}$$

$$A_{IJ} = \frac{\sum_{i=1}^{N_I} [a_i \max(A_{i\bullet})] + \sum_{j=1}^{N_J} \left[a_j \max\left(A_{\bullet j}\right)\right]}{\sum_{i=1}^{N_I} a_i + \sum_{j=1}^{N_J} a_j} \tag{8}$$

In this paper we explore the performance of both the weighted and unweighted statistics. Although we expect that the weighted statistics will produce more reliable results in most situations, calculating polygon areas is time-consuming and greatly increases the expense of the Monte Carlo process. For cases where the unweighted and weighted statistics produce similar results, the unweighted statistics may be used to decrease computation time.

2.2 Randomization procedure

Evaluation of the area overlap statistics occurs through Monte Carlo randomization procedures corresponding to a specific null hypothesis. We assume that the number of polygons may differ across partitions ($N_I \neq N_J$) and that the number of polygons (N_I , N_J) in each partition is given. Hence the randomization procedures replicate these characteristics of the data. For all of our analyses, we used a null spatial model that found an alternative, randomly generated partition for one of the two polygon sets (here, set **I**); the other partition (**J**) was assumed to be fixed. This alternative partition for **I** should possess the same number of polygons as the original (N_I). To accomplish this, we randomly located N_I seed points throughout the geographic study space, and then constructed Voronoi (Thiessen) polygons from the resulting point set. The set of Voronoi polygons served as the alternative partition. The alternative partition was then compared with the second original partition and area overlap statistics were calculated; this process was repeated 500 times for each analysis, resulting in a distribution for each statistic under the null hypothesis. Because larger values of the area overlap statistics represent better overlap, upper p-values were compared with a 0.05 level of significance (α) to determine whether or not the polygon sets overlapped, and lower p-values were similarly compared to determine whether or not the polygon sets displayed avoidance.

Other null hypotheses could be used to assess area overlap statistics; for example, both sets could be randomized each Monte Carlo iteration. Randomizing one set and not the other is an appropriate null hypothesis for cases where **J** (the fixed set) represents a known process and/or constant phenomenon, and for when the alternative model involves **J** somehow giving rise to **I**. For example, see Hall and Maruca (2001), where the alternative model involves vegetation patterns (which are relatively fixed through one breeding season) giving rise to spatial patterns in songbird territories. We call this a conditional randomization. If the alternative model does not include a directional relationship, a null model that randomizes both sets is appropriate, and is called an unconditional randomization. We chose to randomize only one set with our simulated data in order to parallel the Soda Butte analysis (see below). Further, alternative partitions could be found using algorithms other than Voronoi polygon construction on randomly placed seed points; for example, region-growing or adaptive polygon shuffling techniques could be used. The Voronoi algorithm was found to be the most efficient of those we considered.

2.3 Performance with simulated data

Before applying the area overlap statistics to real-world data, we verified them by analyzing 5 simulated scenarios, all but the last of which included two sets of 12 polygons each. As a first step we wished to verify that (1) in an analysis of two sets of perfectly overlapping polygons, the statistics show significant area overlap, and (2) in an analysis of maximally offset polygons, the statistics show significant overlap avoidance. For these situations, we used the two polygon sets in Fig. 3. We conducted three additional analyses with simulated data; in

each, the partitions overlap well but not perfectly (Figs. 4 and 5). These three scenarios represent slightly more realistic data sets, which roughly correspond to levels of area overlap that any set of well-designed area overlap statistics should detect as significant. In the first pair of partitions, the polygons are approximately the same size (Fig. 4A), and in the second pair, the polygon sizes vary (Fig. 4B). In the third pair, the polygons comprise partitions

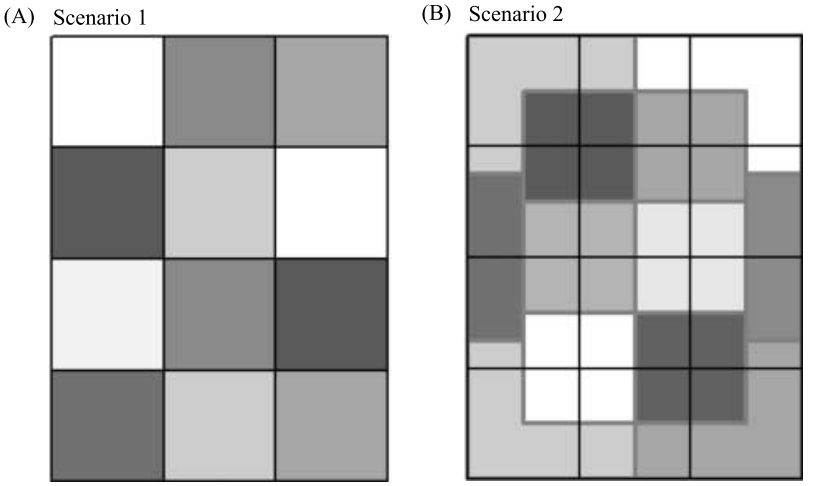

Fig. 3. Simulated scenarios 1 (**A**) and 2 (**B**), each comprised of 12 polygons, used in preliminary tests of area overlap statistics. **A** Sets **I** and **J** overlap perfectly. **B** Polygon outlines for sets **I** (*gray outline*) and **J** (*black outline*) are offset from each other

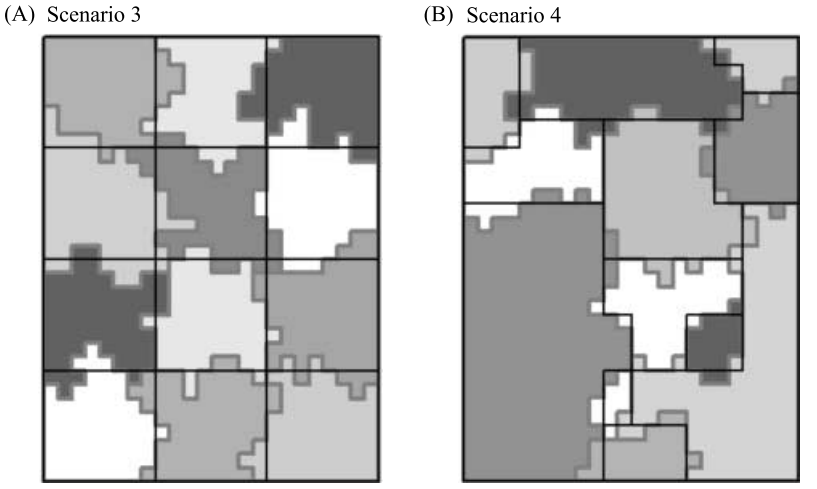

Fig. 4. Simulated scenarios 3 (**A**) and 4 (**B**), each comprised of 12 polygons, used in preliminary tests of area overlap statistics. Each pair of polygon sets **I** (*gray outline*) and **J** (*black outline*) overlap well but not perfectly. **A** Polygons are of similar sizes. **B** Polygons are of different sizes

delineated on different spatial scales (Fig. 5), with one set containing 12 polygons and the second containing 34. Each polygon in the first set of 12 matches perfectly with the union of either 2 or 3 polygons in the second set.

2.4 Performance in an application: Soda Butte Creek

We then used the statistics on high-resolution hyperspectral imagery of the Round Prairie area of Soda Butte Creek in Yellowstone National Park, Wyoming (Fig. 1 in Jacquez et al. 2002 in this special issue). Soda Butte Creek is impacted by heavy metals from past mining activities and a related superfund site (Marcus et al. 2001). Current research focuses on relationships between heavy metal concentrations in stream sediments and stream morphological units, which are also called in-stream habitats (Ladd et al. 1998, Marcus et al. 1996). Mapping the stream morphological units is a tedious process that requires field workers to investigate the entire stream reach and subjectively categorize stream areas as pools, riffles, glides, etc. The accuracy of resulting morphological unit maps is short-lived, in that a single storm event can substantially alter stream morphology. This research would therefore be greatly aided by an automated method for mapping stream morphological units from digital imagery, although limited success has been achieved by such efforts thus far (Wright et al. 2000). In this application, we use area overlap analysis to assess the feasibility of using spatially explicit classification (also called spatially constrained clustering; e.g. Legendre 1987, Legendre and Fortin 1989, Fortin and Drapeau 1995) with high-resolution hyperspectral images to find stream morphological units along a short reach of Soda Butte Creek.

We preprocessed the 1-m resolution image for this analysis by first reducing its nonspatial dimensionality. From 128 possible spectral bands,

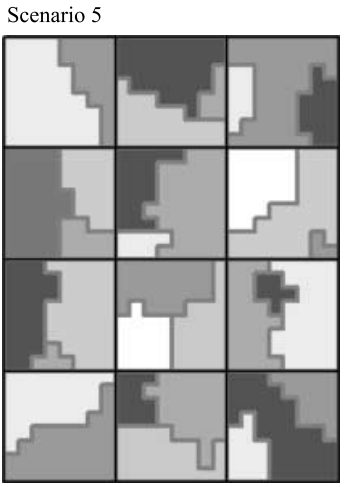

Fig. 5. Simulated scenario 5, used in preliminary tests of area overlap statistics. Polygon sets **I** (*gray outline*, 34 polygons) and **J** (*black outline*, 12 polygons) represent related partitions on different spatial scales

we selected 8 that spanned the spectral range and represented characteristics of the stream important in mapping morphological units. The selected wavelengths (555, 615, 678, 753, 1069, 1230, 1575, and 2143 nm) were sensitive to water surface turbulence and/or water depth. Alternative dimensionality reduction techniques include Principal Components Analysis, but selection of specific bands is reasonable given our prior knowledge of their relevance to known fluvial features. The images were then masked to include only in-stream areas. BoundarySeer, software for geographic boundary analysis, (TerraSeer 2001, www.terraseer.com) was used to perform the spatially explicit classification, which combined an agglomerative, hierarchical method with a redistribution method (k-means). The result is a partition that minimizes within-unit and maximizes between-unit variability within constraints imposed by the desired spatial scale of the mapped stream units. The partition was compared with a digitized map of morphological units delineated by a field mapping crew. During evaluation of the area overlap statistics, the field-mapped polygons were fixed while the automated units were randomized (i.e., we conditioned randomization on the field crew classification), for two reasons. First, the question is whether spatially explicit classification can find units that match the field-mapped units (not vice-versa); and second, the field-mapped units did not exhaustively partition the space (see Fig. 6), and we lacked partitioning algorithms that reproduced the idiosyncrasies arising from the in-field digitizing process.

3 Results

Table 1 shows the results of area overlap analyses involving the five simulated scenarios. In general, the statistics behaved as expected and were

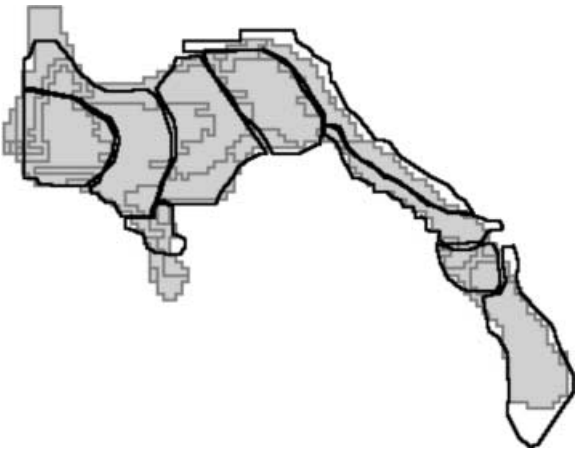

Fig. 6. Stream morphological units for a section of Soda Butte Creek, Wyoming. Stream flow is from top left to bottom right. Polygon set **I** (*gray outline*) represents putative morphological units obtained from spatially explicit classification of high resolution (1 m) imagery. Polygon set **J** (*black outline*) represents morphological unit delineations as determined by a field crew

Table 1. Results of area overlap analyses involving five simulated scenarios (Figs. 3–5). U = statistics unweighted; W = statistics weighted by area. Each P-value calculated from a null distribution comprised of $N = 500$ randomizations. P-values in bold are significant at $\alpha = 0.05$. Significant upper p-values represent overlap, while significant lower p-values represent avoidance

Scenario	Wts.	A_I Value	A_I p_{upper}	A_I p_{lower}	A_J Value	A_J p_{upper}	A_J p_{lower}	A_{IJ} Value	A_{IJ} p_{upper}	A_{IJ} p_{lower}
1. Perfect overlap (Fig. 3A)	U	1.0	**≤0.002**	≥0.998	1.0	**≤0.002**	≥0.998	1.0	**≤0.002**	≥0.998
	W	1.0	**≤0.002**	≥0.998	1.0	**≤0.002**	≥0.998	1.0	**≤0.002**	≥0.998
2. Maximally offset polygons (Fig. 3B)	U	0.299	0.968	**0.034**	0.320	0.992	**0.010**	0.310	0.978	**0.024**
	W	0.310	0.994	**0.008**	0.320	0.994	**0.008**	0.315	0.996	**0.006**
3. 'Good' overlap – polygon areas similar (Fig. 4A)	U	0.743	**≤0.002**	≥0.998	0.743	**≤0.002**	≥0.998	0.743	**≤0.002**	≥0.998
	W	0.742	**≤0.002**	≥0.998	0.743	**≤0.002**	≥0.998	0.743	**≤0.002**	≥0.998
4. 'Good' overlap – polygon areas different (Fig. 4B)	U	0.776	**≤0.002**	≥0.998	0.776	**≤0.002**	≥0.998	0.776	**≤0.002**	≥0.998
	W	0.833	**≤0.002**	≥0.998	0.834	**≤0.002**	≥0.998	0.833	**≤0.002**	≥0.998
5. Overlapping partitions – different spatial scales	U	0.353	**≤0.002**	≥0.998	0.543	**≤0.002**	≥0.998	0.403	**≤0.002**	≥0.998
	W	0.444	**≤0.002**	≥0.998	0.543	**≤0.002**	≥0.998	0.494	**≤0.002**	≥0.998

able to discriminate between the absence and presence of different amounts of overlap. All three scenarios in which the two polygon sets clearly appeared to overlap (scenarios 1, 3, and 4) showed significant overlap unidirectionally (A_I and A_J) and bidirectionally (A_{IJ}; $p_{upper} \leq 0.002$ for each), whether using unweighted or area-weighted averages. In scenario 5, where the two partitions are related but on different spatial scales, the extent of area overlap may be difficult to predict subjectively. Compared with our null spatial model, the unweighted and area-weighted overlap statistics calculated for scenario 5 were significant ($p_{upper} \leq 0.002$ for A_I, A_J, and A_{IJ}). Although polygons in the 34-polygon set are, on average, much smaller than those in the 12-polygon set, their relative arrangements are indicative of a strong relationship between the two partitions. The scenario in which the two polygon sets were maximally offset (scenario 2) showed significant avoidance; that is, the polygon sets overlapped much less than expected by chance, whether the statistics were unweighted ($p_{lower} = 0.034$ for A_I, $p_{lower} = 0.010$ for A_J, and $p_{lower} = 0.024$ for A_{IJ}) or weighted by area ($p_{lower} = 0.008$ for A_I, $p_{lower} = 0.008$ for A_J, and $p_{lower} = 0.006$ for A_{IJ}).

In scenarios 1, 2, and 3, where polygons were similar in area, values for area-weighted overlap statistics were not meaningfully different from the unweighted statistics, as expected. However, values for unweighted and area-weighted statistics were different for the two polygon sets in scenario 4, where polygon areas varied within each set, and for those in scenario 5, where polygon areas varied within set **I** and between the two sets. Recall that overlap is measured by the ratio of area of intersection to area of union, and hence is a relative measure scaled to the range 0–1. Higher values for area-weighted statistics are expected whenever larger polygons overlap more closely than smaller polygons. This phenomenon can readily be seen in the map for scenario 4 (Fig. 4B), where all three statistics displayed higher values when area-weighting was used. In scenario 5, the area-weighted value for A_I was higher than the unweighted value. Because all polygons in set **I** are completely contained within a polygon in set **J**, and because all polygons in set **J** have the same area, the area of union between a polygon in **I** and its best match in **J** is always the same and equal to the area of the **J** polygon; however, the area of intersection is equal to the area of the **I** polygon, and so the measured overlap will be greater for larger polygons in set **I**.

In order to best partition Soda Butte Creek using the 8 selected spectral bands, we first used spatially explicit classification to find a series of partitions, each differing from the others in the number of classes (i.e. polygons) established. We then measured the extent to which each partition 'fit' the data by calculating a goodness-of-fit index (Hall and Maruca 2001), based on the within-class and between-class sum of squares errors. The partition with 30 classes presented the best trade-off between a good fit (accuracy) and a reasonably small number of classes (simplicity), and was therefore selected as the final partition. The nine (9) digitized polygons representing morphological units mapped by the field crew are shown with the 30-polygon automated partition in Fig. 6.

Table 2 shows the results of area overlap analysis of Soda Butte morphological units using unweighted and area-weighted statistics. For both weighted and area-weighted statistics the averages are calculated over all polygons in the segmentation and the order in which these polygons are

Table 2. Results of area overlap analyses of Soda Butte Creek morphological units (Fig. 6). U = statistics unweighted; W = statistics weighted by area. Each P-value calculated from a null distribution comprised of $N = 500$ randomizations. P-values in bold are significant at $\alpha = 0.05$. Significant upper p-values represent overlap, while significant lower p-values represent avoidance

Wts.	A_I Value	A_I p_{upper}	A_I p_{lower}	A_J Value	A_J p_{upper}	A_J p_{lower}	A_{IJ} Value	A_{IJ} p_{upper}	A_{IJ} p_{lower}
U	0.137	0.874	0.128	0.382	**≤0.002**	≥0.998	0.193	0.310	0.692
W	0.369	**≤0.002**	≥0.998	0.412	**≤0.002**	≥0.998	0.390	**≤0.002**	≥0.998

selected makes no difference on the value of the statistic. Using unweighted averages for the overlap statistics, A_J was found to be significant ($p_{\text{upper}} \leq 0.002$), while A_I and A_{IJ} were not. The polygons representing morphological units defined by the field crew (set **J**) appear to overlap well with the units defined by the spatially explicit classification algorithm (set **I**), but not the other way around. Specifically, each field mapped unit overlapped with its best match from the set of original classified units better than when the classified units were randomly located within the study area. The relatively low and statistically insignificant value of A_I may be attributed to the fact that, due to edge effects and slight differences in the study area boundaries between the two sets, 7 (of 30) classified units were small and located on the periphery, and therefore did not overlap any field crew morphological unit (i.e. $A_i = \max(a_{i\bullet}) = 0$ for each). This example demonstrates how area-weighted averages of area overlap can provide more meaningful information about the total overlap between the two polygon sets. Indeed, when area-weighted averages were used, all three overlap statistics were found to be highly significant ($p_{\text{upper}} \leq 0.002$ for A_I, A_J, and A_{IJ}). The stream units with the larger areas in the automated partition overlap very well with the units in the field crew partition, and vice versa.

4 Discussion

The real world is highly complex. Patterns result from space-time processes that operate at different spatial and temporal scales, in overlapping domains, and with varying strength and intensity. These patterns are characterized by anisotropy, non-linearity, non-stationarity, non-ergodicity, and hysteresis. As a result spatial heterogeneity is often large near edges where landscape features intergrade. Edges themselves may be sinuous, crenellate or even fractal, and landscape constituents, such as patches, may change in size with proximity to edges. When assessing the correspondence between patterns defined on real-world systems, we require methods that are founded on assumptions that are consistent with these characteristics. In this paper we have attempted to develop and evaluate such a technique.

In this preliminary analysis, the area overlap statistics correctly identified overlapping sets of polygons in all simulated scenarios where the polygon sets were known *a priori* to overlap, and they indicated avoidance for the polygon sets that were offset from each other. The behavior of these statistics

is certainly promising and warrants further testing to evaluate statistical power, quantify sensitivity to a range of spatial scales and topologies, and to gauge the applicability of this analytic method to real-world questions. Future work should focus on testing the statistics with simulated data generated using process-based models, and on evaluating the sensitivity of the statistics to changes in spatial scale.

The successful implementation of any spatial statistic evaluated via Monte Carlo randomizations depends on the selection of an appropriate null spatial model; therefore, additional analyses involving a range of null models are planned. In general the randomization algorithm must correspond to a reasonable null or neutral physical/biological/ecological model. Hence different randomization procedures may be applicable in different situations, depending on the question being explored and the assumptions applicable under the null model. Alternative null models thus might preserve distributions of polygon area, of polygon shape, number of polygons, and the presence or absence of edge effects. All of these are tractable with the context of our tests.

Alternative null spatial models exist. One alternative to our Voronoi-polygon null model is to devise algorithmically efficient methods for randomizing the original data and regenerating partitions, perhaps using spatially explicit classification or some similar partitioning scheme. Yet, such methods are so computationally intensive as to render them currently infeasible. A second alternative null model would employ random cluster growth under constraints designed to preserve certain characteristics of the system. Because area and shape play so crucial a role in the calculation of the area overlap statistics, an ideal null model would preserve the distribution of unit areas, as well as the general range of unit shapes.

In the analysis of Soda Butte Creek morphological units, the area overlap statistics indicate that there is at least some agreement between stream units defined somewhat subjectively by field crews and stream units defined from high resolution multispectral imagery using spatially explicit classification. Because field crews define morphological units using some of the same stream features to which the spectral reflectances are sensitive, such as water depth and water surface turbulence, statistical agreement between the two maps supports the utility of area overlap analysis. However, if we are to use automated stream units in place of ground-mapped units in other research, we must understand and be willing to accept the differences between partitions that result from using these two very different methods. Perhaps the largest source of discrepancies is that, in defining morphological units, field crews use some criteria that are not necessarily reflected in the spectral data, such as prior knowledge of slope, stream depth, and sediment characteristics; conversely, spectral data may include information that is not necessarily available to the field crews, such as reflectance in the non-visible wavelengths. These differences might be partially overcome by using more or different spectral bands, or by using the first few components of a principal components analysis of the entire spectrum, or by developing classification algorithms that explicitly incorporate adjunct or prior information. But it should be recognized that the differences are highly informative because they can lead to insights regarding the sampling process as well as the classification paradigm being used by humans.

Differences between the classifications can inform how we think about natural variation and the processes embedded in natural systems. Slight differences in the spatial scale of defined units may also lead to disparate partitions; for example, when field crews detect variation on the scale of 1–2 m^2, they will likely consider these areas to be variable regions of larger morphological units rather than as separate units themselves. In general, the field crews are not likely to define very small morphological units, because the human eye and our cognitive systems are sensitive to larger-scale patterns, while units defined by spatially explicit classification are limited in size only by image resolution. In fact, in our example nearly half (13 of 30) the stream units defined using spatially explicit classification were 2 m^2 or less in area. Further, even when unit-defining criteria and spatial scale are comparable in the two methods, map differences may still exist, especially when variation in spectral reflectances does not match human perceptions of natural variability in the stream system.

Are the automated units actually morphological units? Is it important to label them as such? If so, another aid would be to perform spatially explicit supervised classification using spectral signatures for stream units (which do not yet exist). Ultimately, the feasibility of using spatially explicit classification of digital imagery to define stream units rests on the utility of such units in the context of the biological or physical processes under investigation. In their work, Marcus, Ladd, and colleagues (e.g. Marcus et al. 1996, Ladd et al. 1998, Marcus et al. 2001) are interested in understanding the processes that lead to the transport, fate and deposition of heavy metals in streams. Their approach has been to demonstrate the segregation of heavy metals in stream bed sediments according to stream morphological units; however, their ability to detect this relationship relies on the existence of processes whereby metals are differentially deposited and/or retained according to certain stream characteristics (water depth, flow, velocity, etc.) that are used in our definition of stream morphological units (pool, glide, riffle, etc.). It is possible that units defined using spatially explicit classification of digital imagery, though different from human-defined morphological units, may be at least as well-suited to the study of heavy metals in stream bed sediments.

References

Aspinall RJ, Marcus WA, Boardman J (2002) Considerations in collecting, processing, and analyzing high spatial resolution and hyperspectral images in analyses and forecasts. (In this issue). *Journal of Geographical Systems* 4:15–29

Bonham-Carter GF (1994) Geographic Information Systems for Geoscientists. Pergamon, Oxford

Egenhofer MJ, Glasgow J, Günther O, Herring JR, Peuquet DJ (1999) Progress in computation methods for representing geographic concepts. *International Journal of Geographic Information Science* 13:775–796

Fortin M-J (1997) Effects of data types on vegetation boundary delineation. *Canadian Journal of Forest Research* 27:1851–1858

Fortin M-J, Drapeau P (1995) Delineation of ecological boundaries: Comparisons of approaches and significance tests. *Oikos* 72:323–332

Fortin M-J, Drapeau P, Jacquez GM (1996) Quantification of the spatial co-occurrences of ecological boundaries. *Oikos* 77:51–60

Goovaerts P (2002) Geostatistical incorporation of spatial coordinates into supervised classification of hyperspectral data. (In this issue). *Journal of Geographical Systems* 4:99–111

Griffith D (2002) Modeling spatial dependence in high spatial resolution hyperspectral datasets. (In this issue). *Journal of Geographical Systems* 4:43–51

Hall KR, Maruca S (2001) Mapping a forest mosaic: A comparison of vegetation and bird distributions using geographic boundary analysis. *Plant Ecology* 156:105–120

Jacquez GM (1995) The map comparison problem: Tests for the overlap of geographic boundaries. *Statistics in Medicine* 14:2343–2361

Jacquez GM, Maruca S, Fortin M-J (2000) From fields to objects: A review of geographic boundary analysis. *J Geograph Syst* 2:221–241

Jacquez GM, Marcus WA, Aspinall RJ, Greiling DA (2002) Exposure assessment using high spatial resolution hyperspectral (HSRH) imagery. (In this issue). *Journal of Geographical Systems* 4:1–14

Ladd S, Marcus WA, Cherry S (1998) Trace metal segregation within morphologic units. *Environmental Geology and Water Sciences* 36:195–206

Legendre P (1987) Constrained clustering. In: Legendre P, Legendre L (ed) *Developments in Numerical Ecology*, NATO ASI series, Vol G 14. Springer, Berlin, pp 289–307

Legendre P, Fortin M-J (1989) Spatial pattern and ecological analysis. *Vegetatio* 80:107–138

Marcus WA, Ladd S, Crotteau M (1996) Channel morphology and copper concentrations in stream bed sediments. In: Nelson JD et al. (ed) *Tailings and Mine Waste '96*. Balkema Press, Rotterdam, pp 421–430

Marcus WA, Meyer GA, Nimmo DR (2001) Geomorphic control of persistent mine impacts in a Yellowstone Park stream and implications for the recovery of fluvial systems. Geology 29:355–358

Mark DM, Freksa C, Hirtle SC, Llyod R, Tversky B (1999) Cognitive models of geographic space. *International Journal of Geographic Information Science* 13:747–774

Rogerson P (2002) Change detection thresholds for remotely sensed images. (In this issue). *Journal of Geographical Systems* 4:85–97

Sadahiro Y (2000) Exploratory method for analyzing changes in polygon distributions. (White paper.) Center for Spatial Information Science, University of Tokyo, Japan

Sadahiro Y, Umermura M (2000) Spatio-temporal analysis of polygon distributions: event-based approach. CSIS Discussion Paper Series No. 25, Center for Spatial Information Science, University of Tokyo, Japan

Terraseer (2001) BoundarySeer: Software for the analysis of geographic boundaries. TerraSeer, Inc. Ann Arbor, Michigan. www.terraseer.com

Wright A, Marcus WA, Aspinall R (2000) Applications and limitations of using multispectral digital imagery to map geomorphic stream units in a lower order stream. *Geomorphology*, 33:107–120