Landscape Ecology vol. 11 **no.** 3 pp 129–140 (1996)
SPB Academic Publishing bv, Amsterdam

The modifiable areal unit problem and implications for landscape ecology

Dennis E. Jelinski* and Jianguo Wu **
Departmen of Forestry, Fisheries and Wildlife, Institute of Agriculture and Natural Resources, University of Nebraska Lincoln, Lincoln, NE, 68583-0814

Keywords: modifiable areal unit problem, scale, aggregation, zoning systems, spatial analysis, spatial autocorrelation

Abstract

Landscape ecologists often deal with aggregated data and multiscaled spatial phenomena. Recognizing the sensitivity of the results of spatial analyses to the definition of units for which data are collected is critical to characterizing landscapes with minimal bias and avoidance of spurious relationships. We introduce and examine the effect of data aggregation on analysis of landscape structure as exemplified through what has become known, in the statistical and geographical literature, as the *Modifiable Areal Unit Problem* (MAUP). The MAUP applies to two separate, but interrelated, problems with spatial data analysis. The first is the "scale problem", where the same set of areal data is aggregated into several sets of larger areal units, with each combination leading to different data values and inferences. The second aspect of the MAUP is the "zoning problem", where a given set of areal units is recombined into zones that are of the same size but located differently, again resulting in variation in data values and, consequently, different conclusions. We conduct a series of spatial autocorrelation analyses based on NDVI (Normalized Difference Vegetation Index) to demonstrate how the MAUP may affect the results of landscape analysis. We conclude with a discussion of the broader-scale implications for the MAUP in landscape ecology and suggest approaches for dealing with this issue.

Introduction

Increasingly, ecological research is being conducted at larger spatial scales – landscape and regional scales – in large part because of interest in landscape dynamics, biodiversity, and global change (Hall *et al.* 1988; Ross *et al.* 1988; Jelinski *et al.* 1994). Unfortunately, however, much of our knowledge of scale-dependent phenomena derives from the aggregation of area-based information obtained from small areas (less than 1 km^2), represented by even smaller plots (1–30 m^2) (Burke 1991). Thus the choice of the basic areal units (BSUs) for analysis and "scaling up" are often arbitrary (Meentemeyer 1989) or dictated by the resolution of available data (*e.g.*, remotely sensed data). Further, the general absence of rules and methods to effectively deal with multiple-scale spatial phenomena remains a major hiatus to "scaling up" spatial data (Gardner *et al.* 1982; Rastetter *et al.* 1992; Levin 1992, 1993). The problems of aggregation error and inference across scales have been recognized for decades, but have assumed greater importance in landscape ecology with the current focus on analysis of landscape structure, and upscale integration of simulation models and parameter fields (Turner *et al.* 1991; Wiens *et al.* 1993, Rastetter *et al.* 1992; Jelinski *et al.* 1994; Wu and Levin 1994). Additionally, the effects of changing scale on the analysis of spatial pattern

**Address for correspondence:* Department of Forestry, Fisheries and Wildlife, 101 Plant Industry, University of Nebraska-Lincoln, Lincoln, Nebraska 68583-0814, U.S.A.; ** Current address: Department of Life Sciences, Arizona State University West, 4701 West Thunderbird Road, Phoenix, AZ 85069-7100, U.S.A.

and process have been emphasized during the last decade, such as through increasing interest in hierarchy theory (Allen and Starr 1982; O'Neill *et al.* 1986; Wu and Loucks 1995).

Methods to detect the scale of landscape pattern can be traced back to the early work of plant ecologists who recognized and developed methods to handle scale-dependent patterns and processes (*e.g.*, Greig-Smith 1952, 1957, 1979; Kershaw 1957, 1964). Of the work that has been published in the landscape ecology literature on the results from studies of spatially aggregated data, only a few have dealt empirically with the consequences of changing scale. For example, Nellis and Briggs (1989) used textural analysis at three levels of spatial resolution to assess landscape structure of tallgrass prairie grasslands subject to different management regimes. Turner *et al.* (1989) studied the scale effects in landscape pattern analysis, using indices measuring diversity, dominance, and contagion. Data from USGS land use maps and computer-generated random maps showed the existence of thresholds in spatial patterns. A comprehensive treatment of statistical methods for scale-detection using landscape data was provided by Turner *et al.* (1991). Notwithstanding these treatments of aggregation problems, the effect of different zoning systems used in the aggregation process on data values and inferences has received particularly little attention in ecological research.

Though virtually unknown in the landscape ecology literature, the most comprehensive treatment of the sensitivity of analytical results to the definition of data collection units is found in the statistical and geographical literature, where it is known as "the Modifiable Areal Unit Problem" (MAUP) (Openshaw and Taylor 1979, 1981; Openshaw 1984). The modifiable areal unit problem arises from the fact that areal units are usually arbitrarily determined and "modifiable", in the sense that they can be aggregated to form units of different sizes or spatial arrangements. Thus the MAUP has two related but distinctive components: the scale problem and the zoning (or aggregation) problem (Openshaw and Taylor 1979, Openshaw 1984). The scale problem is "the variation in results that may be obtained when the same areal data are combined into sets of increasingly larger areal units of analysis". The zoning problem, in contrast, is "any variations in results due to alternative units of analysis where n, the number of units, is constant" (Openshaw and Taylor 1979). For any specified number of zones, there are many ways of defining the boundaries of these zones.

In the contrived example shown in Figure 1(*a-c*), one can see the effects of aggregating areal data from neighbouring zones as explified through the MAUP. While the mean value does not change, the variance declines with increasing aggregation. As a consequence of this smoothing effect, information on spatial heterogeneity in the landscape (patchiness) is lost or distorted. In Figure 1(*d-f*), units (pixels) have been aggregated into zones with varying orientations of the cardinal directions. For parts d and *e*, there is not much change in window mean; however, variance changes substantially as a function of location. By comparing parts c, *e* and *f* one can see that even when the number of zones is held constant ($N = 4$) the mean and variance is affected. Moreover, a comparison of parts *b* and d show a change in variance when the orientation is altered but the size of the units remains fixed.

Openshaw and Taylor (1977) studied the effects of the MAUP through three related experiments under different spatial and statistical conditions. The basic areal units in the data set were the 99 counties in the state of Iowa. By correlating the percentage of elderly voters with Republican voters in Iowa, they showed that if the 99 counties making up the state were grouped together into fewer larger districts, and all possible combinations of the larger-scale districts were considered, correlations ranging from +0.979 to –0.811 could be produced by varying the scale and zoning strategies. Similar results were found in several earlier studies of aggregation effects (*e.g.*, Gehlke and Biehl 1934; Yule and Kendall 1950; Robinson 1950). The MAUP also carries implications for multivariate statistical analyses and spatial interaction models. Fotheringham and Wong (1991) showed that model calibration was sensitive to variations in scale and zoning systems, leading to highly unreliable results in multivariate analyses. It is noteworthy that all the aforementioned examples are based on geographical data where the basic areal units usually vary in size and shape corresponding to administrative or political boundaries.

The spatial association or mosaic patterning

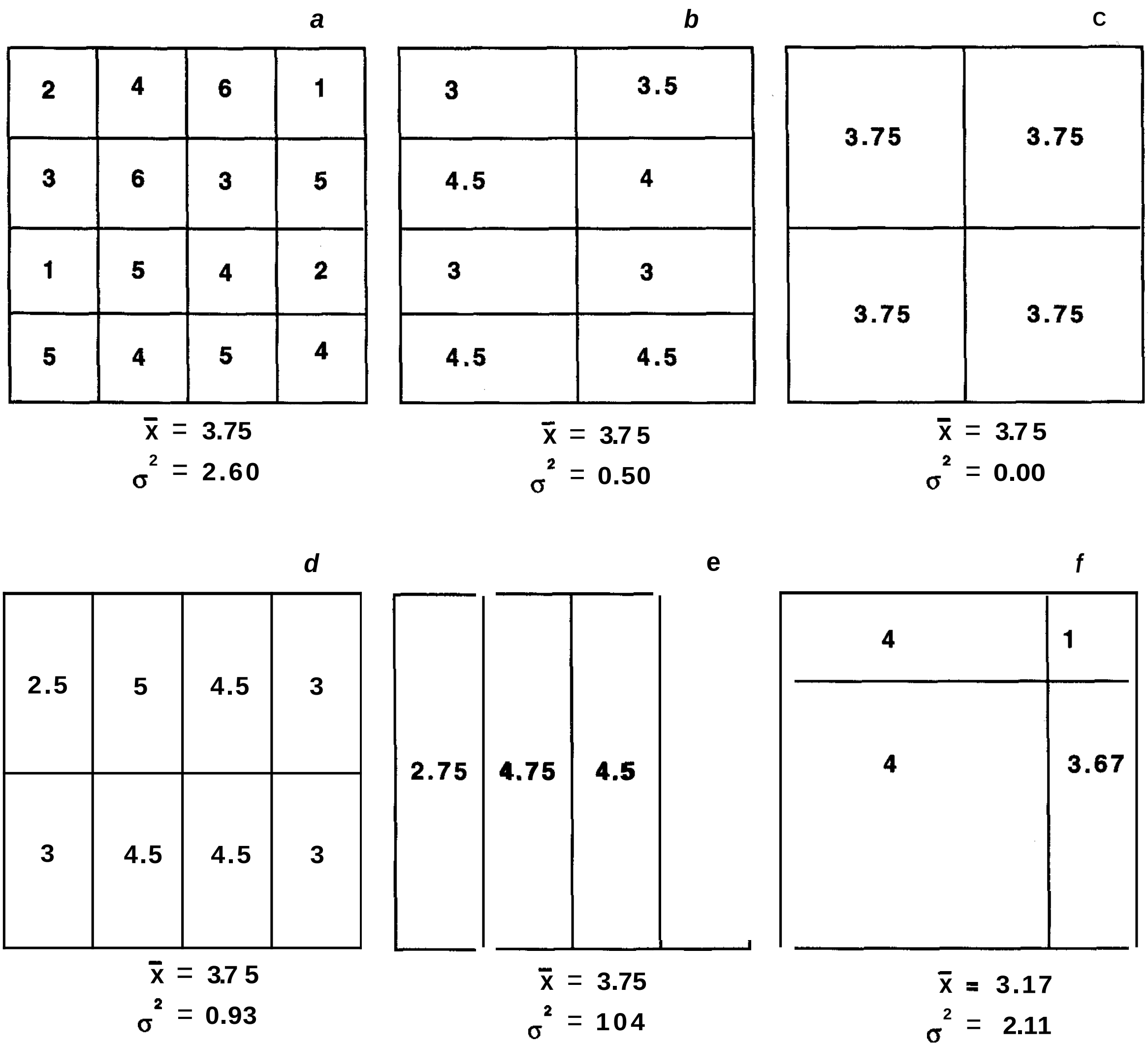

Fig. 1. Contrived example showing the two interrelated aspects of the modifiable areal unit problem.
a–c. Effects of areal aggregation
d–f. Effects of various zoning systems

among species in grassland communities (*e.g.*, Greig-Smith 1983) presents another illustration of the MAUP. Typically, an ecologist would lay out a rectangular grid that is composed of a large number of continuous equal-sized square cells, over which frequency or density data of the species under study are collected. The size of the grid cell and the extent of the entire grid are usually determined by convention (*e.g.* 1×1 m^2 for herbaceous vegetation). The degree of association among these species, as measured by some statistic, changes when the data are aggregated from the basic areal unit (single grid cell) to larger blocks (see Greig-Smith 1983). This type of variation in a measure with the level of aggregation is typically the scale problem. In contrast, there are a large number of ways to arrange the areal units or to define the zones (blocks) within each level of aggregation (especially at lower scales in this case). The variation in a measure (*e.g.*, correlation coefficient, diversity index) caused by alternative blocking strategies is the zoning problem. This aspect of sampling and patch delineation has received particularly scant attention.

Recognition of patches (aggregates of individuals and communities) in particular is fundamental to understanding landscape structure (Watt 1947; Forman and Godron 1986; Urban *et al.* 1987; Wu and Levin 1994). Wherever areal data are used and aggregated to define patch structure, the modifiable areal unit problem may occur. It is critical to understand how the results of landscape analysis may be affected by both scale and zoning problems. Moreover, we believe that the MAUP in general has widespread implications for landscape ecological studies. In the following section, we shall illustrate the essence of the MAUP, through an example with remote sensing data where an $N \times N$ pixel grid is used. The effects of aggregation in the context of the MAUP are examined in terms of spatial autocorrelation. Spatial autocorrelation permits tests of hypotheses regarding spatial patterns (Cliff and Ord 1973) by considering if the presence of a factor in a place makes its presence in neighboring places more likely (positive; aggregation) or less likely (negative; segregation). Thus the test both describes the structure of a spatial pattern and is also capable of detecting the presence of directional components (clinal trends) at various scales (Legendre and Fortin 1989). We recognize that a thorough analysis of patch size and pattern should include more than one method of spatial analysis (Cullinan and Thomas 1992; O'Neill *et al.* 1991). The main purpose here is to illustrate the MAUP, rather than present the results of a detailed analysis of landscape structure. We also discuss several ways of constructively dealing with the modifiable areal unit problem.

Methods

Landscape data

To characterize the structure of the landscape we calculated NDVI (Normalized Difference Vegetation Index) from three Landsat Thematic Mapper (TM) scenes. The first is of the Boreal Forest of north central Manitoba, Canada (taken September, 1990). The study area has gentle relief, with a few lakes but many wetlands and treed bogs, and has an upland that is covered primarily with black spruce *(Picea mariana)*, birch *(Betula balsamifera)* and jack pine (*Pinus banksiana*). Periglacial features include peat plateaus and boulder fields. Fire is a common agent of disturbance. The second landscape is an area of intensive row-crop agriculture (primarily irrigated corn) near York, Nebraska (taken August 1992). There is very little topographic relief in the area. The final scene is a native mixed-grass prairie near Mullen, Nebraska (taken August 1992). The topography is gently rolling. The vegetation is a mixture of some grasses and forbs of the tallgrass and shortgrass prairies, with tallgrass species in moister sites and shortgrass species in drier places. There is little cropland in the area.

The original data are at a nominal resolution of 30 m^2, from which we created a landscape with linear dimensions of 300 x 300 pixels. NDVI data can be aggregated to define patches of vegetation and other landscape configurations (e.g. corridors). The NDVI expresses the difference between the incident radiation reflected by photosynthetically-active pigments in green leaves, and that portion reflected in the near-infrared part of the spectrum. This metric is based on the differences in the contribution to the total reflectance in solar wavelengths (*i.e.*, the albedo) of the visible (0.52-0.60 pm) and near-infrared (0.76-0.9 pm) portions of the spectrum. Thus, for growing green vegetation, the reflectance in the near-IR is greater than that in the visible; a condition which is in sharp contrast to the situation typical of bare soils, rock, and snowcover. It is this unique live vegetation signature that makes NDVI an important measure. The NDVI is a difference ratio of the radiances computed based from the following formula:

$$\text{NDVI} = \frac{\text{B4} - \text{B3}}{\text{B4} + \text{B3}}$$

where

B4 = brightness value from infrared band 4 of TM
B3 = brightness value from red band 3 of TM

Thus the NDVI is a bounded ratio that varies between -1.0 and $+1.0$, with only actively growing vegetation having positive values (typically between 0.1 and 0.6). Although there is some sensitivity of the NDVI to parameters such as the sun-sensor-target geometry and backscatter from atmospheric aerosols, the dominant signal appears to be some combination of surface conditions (Box *et al.* 1989).

Scale problem

To investigate the effects of scale (*i.e.* grain size) of the landscape, we aggregated groups of n adjacent pixels into a single data unit. Pixels were aggregated in arrays from the finest aggregation 1 x 1 (n = 1 original pixels per aggregate unit, *i.e.*, Basic Spatial Unit (BSU)) through the coarsest aggregate (n = 225 pixels per aggregate unit). Thus there were 90,000 replicates for the 1 x 1 matrix and 400 replicates for the 15 x 15 matrix. For both these analyses we ran moving windows (picture elements in a neighbourhood of an image data set) of varying size and orientation across the landscape.

Zoning effects

Two systematic aggregation procedures were developed to investigate the zoning effects at separate scales. The criterion for both was equal numbers of pixels per zone. The aggregations differ in the orientation of zones. Zones were first run in a east-west and then north-south direction. That is, for example, a 1 x 100 would indicate a zone of the east/west direction and a 100 x 1 would indicate a zone of the south/north direction.

1) Zoning System at Small Scale – Five alternative zones at the 16 BSU scale had the following dimensions (windows) for equal area zones:

1 × 16, 2 × 8, 4 × 4, 8 × 2 and 16 × 1

2) Zoning Systems at Large Scale – Nine zoning alternatives at the 100 BSU scale had the following dimensions:

1x 100, 2 x 50, 4 x 25, 5 x 20, 10 x 10, 20 x 5, 25 x 4, 50 x 2 and 100 x 1

Statistical analysis

Spatial autocorrelation describes the degree of spatial clustering, that is, the degree to which values at one locality are determined in part by values at neighboring locations. The two most commonly used measures of spatial autocorrelation in geographically-referenced ecological data are Moran's I statistic and Geary's c statistic (Legendre and Fortin 1989). To detect spatial autocorrelation, a matrix is constructed that represents the site relationships in geographic space. Moran's I is based on a cross-product computation of centered data:

$$I = \left(n \,/ \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} c_{ij} \right) \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} c_{ij}(x_i - \bar{x})\,(x_j - \bar{x}) / \sum_{i=1}^{i=n} (x_i - \bar{x})^2$$

where n is the total number of areal units over the entire landscape, x_i and x_j are values of areal units i and j, and $\bar{x}$ is the mean of all areal units, and c_{ij} denotes the connectivity between areal units i and j, taking a value of 1 if areal units i and j are adjacent and 0 otherwise. Moran's I may vary from positive to negative, depending on the values of the variable for the locations connected by a particular adjacency matrix. I approaches 1 when adjacent localities have similar values (positive autocorrelation) and negative values when adjacent localities have dissimilar values (negative autocorrelation).

We also calculated Geary's c as another measure of spatial autocorrelation. Geary's c is a distance-type coefficient derived by summing squared differences between adjacent pairs of values. It is thus more sensitive to absolute differences between paired localities. Geary's c is calculated as

$$c = \left((n-1) \,/ (2) \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} c_{ij} \right) \sum_{i=1}^{j=n} \sum_{j=1}^{i=n} c_{ij}(x_i - x_j)^2 \,/ \sum_{i=1}^{i=n} (x_i - \bar{x})^2$$

in the same notation as previously. For Geary's c, positive spatial autocorrelation is indicated by a value smaller than its expected mean of 1.

Results and discussion

The global working null hypothesis is that there is no spatial autocorrelation for a broad range of aggregation and zoning schemes. While our results do not include any formal testing using a p value for any individual hypothesis, it is clear that spatial

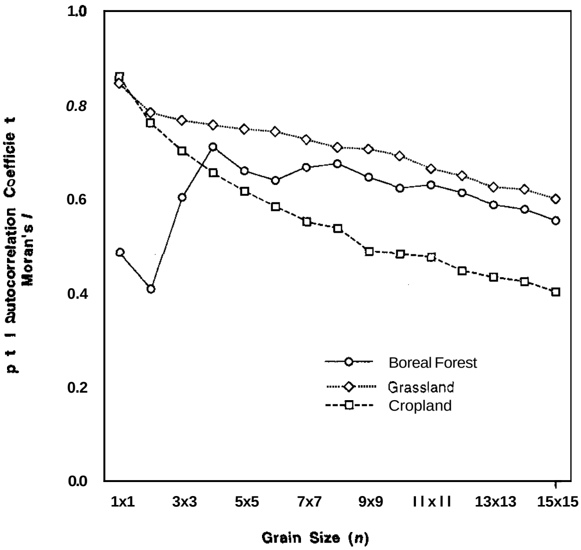

Fig. 2. Effects of data aggregation on spatial autocorrelation resulting from aggregation procedures for a 300 x 300 matrix of 30m^2 pixels of NDVI for Boreal Forest, Grassland and Cropland landscapes.

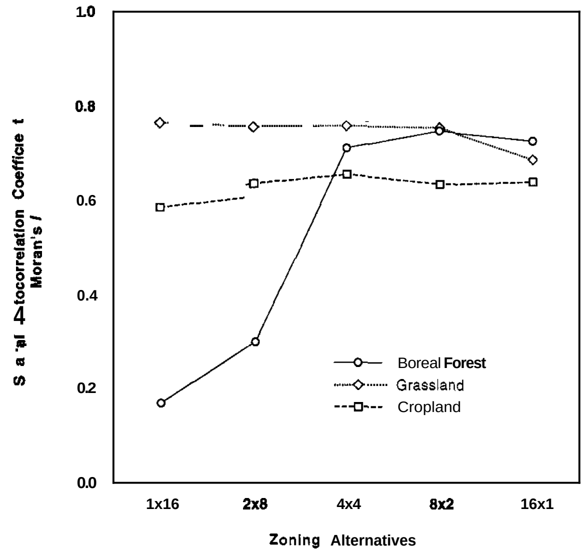

Fig. 3. Effects of selected zone systems on spatial autocorrelation at the 16 BSU scale for a 300 x 300 matrix of 30m^2 pixels of NDVI for Boreal Forest, Grassland and Cropland landscapes.

autocorrelation is not uniform across the landscapes under different aggregation and zoning treatments. Specifically, aggregation effects are captured in Figure 2, which describes the spatial autocorrelation of the various NDVI aggregate units. The first value of spatial autocorrelation for the Boreal Forest site (1 x 1 – 30 m^2 pixel) is positive and then decreases, suggesting that patch size is smaller than the 2 x 2 aggregate (*i.e.*, high patchiness exists at that scale). However, for 3 x 3 to 4 x 4 windows at this site, spatial autocorrelation trends upwards to a peak ($I > 0.7$), which presumably, corresponds to increased homogeneity in landscape structure. At larger zone size, spatial autocorrelation gradually tapers off, though it remains relatively high (*c.* 0.6). In contrast, for the Grassland and Cropland sites, there is a high degree of spatial autocorrelation at the outset. At both sites, autocorrelation then tapers off, but at very different rates, with smaller degree of autocorrelation, *c.* 0.42, for the Cropland compared to *c.* 0.6 for the Grassland sites at the largest grain size 15 x 15). The conclusion is, therefore, that autocorrelation changes with scale and hence there is evidence of a scale effect related to the MAUP among the various aggregations. Geary's *c* mirrored this pattern and thus these results are not reported here.

Figure 3 illustrates the effect of zoning alternatives. For the Boreal Forest site, the pattern of the indices implies directional patchiness for both large and small scale variation in zones, and thus a strong zoning effect. For example, an examination of Figure 3 shows that a window size of 1 x 16 yields a spatial autocorrelation value in sharp contrast to a window with the dimensions of 16 x 1. The same holds true for the case of a 2 x 8 versus 8 x 2. Most importantly, the pattern is different in Figure 4 where spatial autocorrelation is low at the 1 x 100 and 2 x 50 zones, but increases rapidly at the 4 x 25. Autocorrelation remains moderately high and drops off slightly to 0.6. Thus there is more similarity in structure for zones configured in the range from 4 x 25 through 100 x 1 than for the 1 x 100 and 2 x 50 windows. The cause of orientation in landscape structure may be linked to surface geology, which has been suggested to affect forest community development (Elliot-Fisk, 1988). Superimposed on this cause are vegetation and its disturbance history. The combination of these factors affect the distribution and configuration of patch types. The effects of zoning for the Grassland and Cropland site are in sharp contrast to that for the Boreal Forest site. There is no zoning effect at the 1 x 16 through 16 x 1 zoning alternatives (Figure 4). Similarly, there is little change for the Cropland site for all alternative except the 100 x 1, which makes intuitive sense in that cropped areas are

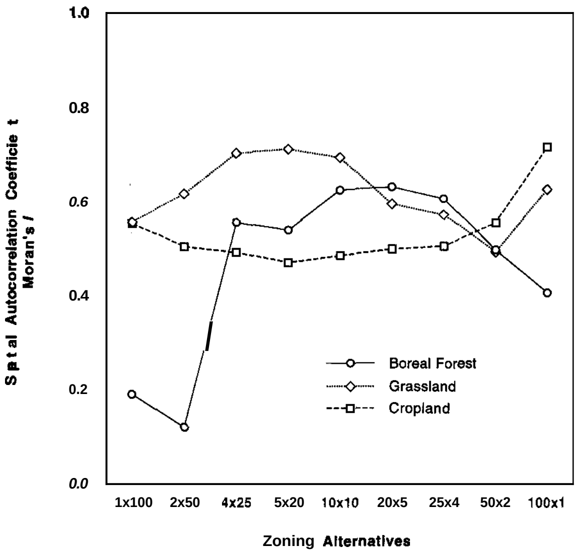

Fig. 4. Effects **of** selected zone systems on spatial autocorrelation at the 100 BSU scale for **a** 300 x 300 matrix of $30m^2$ pixels of NDVI for Boreal Forest, Grassland and Cropland landscapes.

comparatively free of natural disturbance. Zoning alternatives do, however, seem to differentially affect grasslands. Explanations for this may be related to geomorphic structure of the landscape and (or) disturbance such as fire. In sum, the dramatic difference in spatial autocorrelation between the Boreal Forest landscape and the Grassland and Crop landscapes at smaller scales may reflect highly conspicuous patchiness at these scales in the boreal landscape, while the other two are relatively homogeneous. Further analysis of the underlying processes are needed for a fuller explanation of the patterns shown here.

Implications of MAUP for landscape ecology

We have shown above the nature and extent of the interrelated scale and zoning effects inherent in the Modifiable Areal Unit Problem. Here we discuss the ramifications of MAUP more generally in landscape ecology, and then suggest possible ways forward. First, MAUP has implications for the applications of methods for spatial analysis such as Greig-Smith's (1952, 1983) agglomerative contiguous quadrat (blocking) method. In this approach pairs of adjacent quadrates are successively combined into blocks of two through a hierarchy of block sizes (*e.g.*, 2, 4, 8, 16, ... *n*). Means square variances are then plotted against block size. The peaks or clustering revealed in the mean square variance versus block size plot are believed to indicate the scales at which the patterns of population distributions take place (see Greig-Smith 1952, 1983; Kershaw 1957, 1964; Wu 1992). Other workers (Usher 1969; Errington 1973; Cressie 1993) have noted several types of problems that confound analyses with this type of approach. Similarly. it has been noted that the starting position for the blocks affects the outcome (Errington 1973; Upton and Fingleton 1985). This is an indicator of the existence of the MAUP with this method. However, the zoning or block configuration problem was essentially ignored in the original quadrat blocking method.

The MAUP may also affect the results of spatial simulation models when aggregation is involved in the modelling process. There have been studies of spatial interaction modelling which showed that both scale and zoning changes affected the model goodness-of-the-fit and parameter estimates (*e.g.*, Openshaw 1977; Amrhein and Flowerdew 1989; Putman and Chung 1989). Grid-based modelling approaches have been used extensively in landscape ecology in recently years (see Turner and Gardner 1991 for examples). However, little attention has been paid to how the choice of the grid-cell size and aggregation procedures would affect the results and interpretations of the simulation models. There is no theoretical and empirical evidence precluding the existence of the MAUP in such approaches.

Remote sensing observations are playing an increasingly large role in the study of landscape and regional change (Hall *et al.* **1988,** Baker 1989). Low-cost satellite data is becoming more readily available, and many scientists interested in processes and patterns amenable to remote sensing are making use of this data (Roughgarden *et al.* 1991). For example, the United States Geological Survey EROS Data Center recently released low cost five-channel, NDVI (Normalized Difference Vegetation Index) AVHRR (Advanced Very High Resolution Radiometer) data set for the conterminous U.S. (Loveland *et al.* 1991). **A** similar but coarser Global Ecosystems Database was recently released by the National Geophysical Data Center (NGDC) and the U.S. Environmental Protection Agency (EPA) (but see Williams and Jelinski 1995 for description of problems with initial data release). In remote sensing, the modifiable units are the pixels of the image, which represent the fundamental level of spatial resolution that is essentially determined by the capabilities of the sensor and corresponding technology. When different sensors (*e.g.*, Landsat Thematic Mapper (TM), Landsat Multispectral Scanner (MSS), or Systeme Pour l'Observation de la Terre (SPOT)) are used or when pixels are aggregated, these areal units are "modified". Given the size of, say, the Loveland *et al.* (1991) database (2889 rows by 4587 columns), and in view of rapidly increasing use of GIS technology (Jelinski *et al.* 1994), it is highly probable that users of this database will be developing algorithms for different levels of aggregation, different window sizes and different zoning systems. Substantive errors may be introduced during such aggregation procedures if close attention is not paid to the rules for aggregation as shown herein. Several other empirical studies have demonstrated that the results of analyses for the same area may vary due to the variation in spatial resolution of the imagery (see Johnson and Howarth 1987, Woodcock and Strahler 1987, Townshend and Justice 1990).

Approaches for dealing with the MAUP

Given the influence of the modifiable areal unit problem in spatial studies, the development of solutions to it is critically important. A number of ways to understand the impact of the modifiable areal unit problem on spatial analysis have been suggested in the geographical literature (see Openshaw and Taylor 1981; Openshaw 1984; Fotheringham 1989). We examine these alternatives in relation to landscape ecology.

1) **A** *basic entity approach.* This approach advocates the identification of individual entities that, in this case, are ecologically meaningful and not modifiable, and to perform analysis directly on them (Fotheringham 1989). This approach would completely avoid the MAUP, because the MAUP essentially stems from the modifiable nature of areal units and aggregation. Openshaw (1984) maintained that "The usefulness of many forms of spatial study, quantitative or otherwise, depends on the nature and intrinsic meaningfulness of the objects that are under study". In view of this, individual-based (*e.g.*, Pacala and Silander 1985; DeAngelis and Gross 1992) and patch-based (Wu and Levin 1994) approaches in spatial modelling are less susceptible to, if not free from, the modifiable areal unit problem. However, there are difficulties with this basic entity approach. First, it is not always possible to identify what the basic entities are (except for units such as trees, animals, or gopher mounds). Measures such as density, fluxes and the like typically involve integration of spatially distributed data (Hall *et al.* 1988). What are the basic entities for them? Also, there are situations where studies become impossible or impractical because too much detail leads to overwhelming complexity in the data collection, analysis, or modelling of the system under investigation, though the basic entities are identifiable. For example, it is impractical and of little utility to incorporate all individual plants explicitly in an ecosystem study of nutrient cycling at a watershed or regional scale.

*2) **An** optimal zoning approach.* In a series of papers, Openshaw (1977, 1984; Openshaw and Taylor 1979, 1981) suggested a "new paradigm" in spatial analysis of areal data, which involves the derivation of an "optimal" zoning system, *i.e.*, a system that maximizes interzonal variation and minimizes intrazonal variation. Although this optimal zoning approach can avoid the variations in results of analysis caused by the MAUP, optimality is subjective in both definition and operation. In particular, the definition of optimality will change with types of problems under study and statistical methods used. In addition, a zoning system that is optimal for one variable may not be optimal for some other variable. For example, in their study of the MAUP in multivariate statistical analysis, Fotheringham and Wong (1991) concluded that an optimal zoning system that minimizes spatial autocorrelation of all possible combinations of variables is not possible.

3) **A** *sensitivity analysis approach.* Instead of avoiding the problem, an alternative approach to the MAUP is to get a sense of its scope and magni-

tude. By performing a series of sensitivity analysis, one can address the following two questions: What variables are sensitive to the variations in scale and zoning configuration? How sensitive? Conclusions from such studies should, therefore, be scale specific and zoning-system explicit. This approach will advance our understanding of both the phenomenon under investigation and the MAUP in general. However, when the number of variables, the number of scales (levels of aggregation), and the number of zoning alternatives are large, the amount of work to perform a complete sensitivity analysis may become impractical (*e.g.*, the computing demand could easily exceed computational resources available to the researcher).

4) *Development of new methods of analysis.* Another solution is to abandon traditional statistical methods which have been found sensitive to the MAUP (Tobler 1989). It has been suggested that more emphasis should be put on visualization of data than on statistical analysis; for example, representing the data visually over a range of scales is preferable to performing a certain type of spatial analysis at only one level (Openshaw *et al.* 1987, 1988; also see Fotheringham 1989). On the other hand, there has been a call for developing and using spatial analysis methods which are independent of spatial coordinates that are used for collecting and analyzing the data – "frame independent spatial analysis" (Tobler 1989).

Tobler (1989) asserted that all methods whose results depend on areal units should be discarded *a priori*, and that only those techniques independent of areal units should be used. Though it is necessary to develop and use improved spatial statistics, Tobler (1989) underestimated the scope and magnitude of the modifiable areal unit problem, and deemphasized the insights that can be gained from studying the sensitivity of different methods to changes in scale and aggregation (*e.g.*, Openshaw 1984; Turner *et al.* 1989; Fotheringham and Wong 1991). In addition, Openshaw and Taylor (1981) argued that "a context-free approach is contrary to geographical common sense irrespective of what other virtues it may have". They concluded that the right solution to the MAUP should be geographical rather than purely statistical or mathematical.

5) *Emphasis of spatial analysis on the rates of change.* Fotheringham (1989) suggested to shift the emphasis of spatial analysis towards relationships that focus on rates of change by asking the following questions: Can we acquire information on the rate of change in variables and relationships of interest with respect to scale? Do some variables and relationships show erratic fluctuations with scale changes while others do not? Fotheringham (1989) suggested the use of fractal dimension as a scale-independent measure of a spatial distribution or a spatial relationship. The approach is similar to sensitivity analysis in that both focus on the examination of across-scale spatial characteristics of the variables or relationships under consideration.

Fractal dimension remains constant only within the range of spatial scale where self-similarity exists. Even though fractal dimensions are scale-independent over some range of scales, the pattern and process in the real world may still change with scale (Wiens and Milne 1989). To make such an approach effective and feasible, a hierarchical perspective is critically important. We argue that a more comprehensive and effective approach to the MAUP is to develop a conceptual framework based on hierarchical theory (Allen and Starr 1982; O'Neill *et al.* 1986, Urban *et al.* 1987) as opposed to a specific technique. Most if not all systems in nature are hierarchically structured. Certain patterns and processes occur on certain "characteristic" or "threshold" scales. Of course, the so-called characteristic scales are not points in the spatial dimension, but more appropriately finite ranges or domains of scales (see Milne 1988 and Wiens and Milne 1989). When the scale of the observational window matches the characteristic scale of the phenomenon of interest, we will see it; otherwise we miss it. These arguments form the premise of a hierarchical approach to the modifiable areal unit problem. A suggested procedure to deal with the MAUP is simply thus: first to identify the characteristic scales using methods such as spatial autocorrelation, semivariograms, fractal analysis, and spectral analysis, and then to focus the study on these scales.

Conclusions

From the discussion in the previous sections, the presence of the modifiable areal unit problem is not only real, but probably also ubiquitous in many spatial investigations. Openshaw (1984) regarded the MAUP as a fundamental geographical problem inherent in all studies of spatially aggregated data because the results of such studies are always affected by the areal units used. More recently, Fotheringham and Rogerson (1993) listed the MAUP as the first among the eight impediments that arise in spatial analysis. Much greater attention must be paid to this issue in order to advance our understanding and improve the predictability of spatiotemporal pattern and process in nature. This is especially true considering the release of easily accessible remotely-sensed data sets, the rapidly increasing use of GIS, and the universally recognized need for scaling up in ecological studies from landscape to global levels.

It has long been noted that aggregation of spatial data involves a smoothing or filtering effect. Specifically, "aggregating smaller areal units into regions filters out the harmonics whose wavelengths are smaller than the size of the regions" (Tobler 1989). For simple traditional statistical analyses (*e.g.*, correlation analysis, linear regression), such changes can be theoretically expected and thus are relatively well understood (Tobler 1989; Fotheringham and Wong 1991). In general, the zoning problem is much less well understood even for simple statistical analyses. The impact of the MAUP on multivariate analyses has been demonstrated to be far more complex and unpredictable, and yet little work has been done to characterize those effects. Therefore, much more research is needed to unravel the potential impacts of the MAUP in this area.

From a hierarchical stand point of view, the MAUP is not really a "problem", ***per se;*** rather, it may reflect the "nature" of the real systems that are hierarchically structured. It does not, therefore, present any real impediment to understanding spatial phenomena if recognized and dealt with explicitly. On the contrary, it carries critical information we need to understand the structure, function and dynamics of the complex systems in real world. Ecological theories and models have been profoundly influenced by the balance of nature notion and the assumptions of homogeneity and determinism, failing to explicitly recognize the effects of spatiotemporal heterogeneity, scale and multiplicity of ecological systems (DeAngelis and Waterhouse 1987; Jelinski unpubl. ms, Wu and Loucks 1995). The study of the modifiable areal unit problem should provide insight for improved aggregation methods, process-oriented models, and paradigms in ecology.

Acknowledgements

We thank Ye Qi for his generous assistance in computer programming and M. Palecki for critical review of the manuscript. DEJ was supported by National Aeronautics Space Administration (NASA) under grant NAG 5-2331, the National Center for Geographic Information and Analysis (NCGIA) and Mc-Intire-Stennis Funds to the University of Nebraska-Lincoln. This paper is published as University of Nebraska, Agricultural Research Division, Journal Series 11204.

References

Allen, T.F.H. and Starr, T.B. 1982. Hierarchy. University of Chicago Press, Chicago.

Amrhein, C. and Flowerdew, R. 1989. The effect of data aggregation on a Poisson regression model of Canadian migration. *In*: Accuracy of Spatial Databases. pp. 229–238. Edited by M.F. Goodchild and **S.** Gopal, Taylor and Francis, London.

Box, E., Holben, B.N. and Kalb, V. 1989. Accuracy of the AVHRR vegetation index as a predictor of biomass, primary productivity and net CO, flux. Vegetation 80: 71–89.

Burke, I.C., Kittel, T.G.F., Lauenroth, W.K., Snook, P., Yonker, C.M. and Parton, W.J. 1991. Regional analysis of the Great Plains. BioScience 25: 685–692.

Cliff, A.D. and Ord, J.K. 1973. Spatial Autocorrelation. Pion, London.

Cressie, N.A.C. 1993. Statistics for Spatial Data. J. Wiley, New York.

Cullinan, V.I. and Thomas, J.M. 1992. A comparison of quantitative methods for examining landscape pattern and scale. Landscape Ecology 7: 211.

DeAngelis, D.L. and Gross, L.J. (eds.) 1992. Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems. Chapman & Hall, New York.

DeAngelis, D.L. and Waterhouse, J.C. 1987. Equilibrium and nonequilibrium concepts in ecological models. Ecological Monographs 57: 1–21.

Elliott-Fisk, D.L. The Boreal Forest. In North American Terrestrial Vegetation, pp. 33–62. Edited by M.G. Barbour and W.D. Billings. Cambridge University Press, Cambridge.

Errington, J.C. 1973. The effect of regular and random distributions on the analysis of pattern. Journal of Ecology 61: 99–105.

Forman, R.T.T. and Godron, M. 1986. Landscape Ecology. John Riley & Sons, New York.

Fothedngham, A.S. 1989. Scale–independent spatial analysis. *In:* Accuracy of Spatial Databases, pp. 221–228. Edited by M.F. Goodchild and S. Gopal. Taylor and Francis, London.

Fotheringham, AS. and Rogerson, P.A. 1993. GIS and spatial analytical problems. International Journal of Geographic Information Systems 7: 3–19.

Fotheringham, A.S. and Wong, D.W.S. 1991. The modifiable areal unit problem in statistical analysis. Environment and Planning A 23: 1025–1044.

Gardner, R.H., Cale, W.G. and O'Neill, R.V. 1982. Robust analysis of aggregation error. Ecology 63: 1771–1779.

Greig-Smith, P. 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. Annals of Botany, New Series 16: 293–316.

Greig-Smith, P. 1957. Quantitative Plant Ecology. Academic Press, New York.

Greig-Smith, P. 1979. Pattern in vegetation. Journal of Ecology. 67: 755–779.

Greig-Smith, P. 1983. Quantitative Plant Ecology, 3rd ed. University of California Press, Berkeley,

Hall, F.G., Strebel, D.E. and Sellers, P.J. 1988. Linking knowledge among spatial and temporal scales: Vegetation, atmosphere, climate and remote sensing. Landscape Ecology 2: 3–22.

Jelinski, D.E., Goodchild, M. and Steyaert, L. 1994. Multiple roles for GIS in global change research: towards a research agenda. *In*: Environmental Information Management and Analysis: Ecosystem to Global Scales. pp. 41–56. Edited by W.K. Michener, J.W. Brunt and S.G. Stafford, Taylor & Francis, London.

Johnson, D.D. and Howarth, P.J. 1987. The effects of spatial resolution on land cover/land use theme extraction from airborne digital data. Canadian Journal of Remote Sensing 13: 68–74.

Kershaw, K.A. 1957. The use of cover and frequency in the detection of pattern in plant communities. Ecology 38: 291–299.

Kershaw, K.A. 1964. Quantitative and Dynamic Ecology. Edward Arnold, London.

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

Levin, S.A. 1992. The problem of pattern and scale in ecology. Ecology 73: 1943–1967.

Levin, S.A. 1993. Concepts of scale at the local level. *In:* Scaling Physiological Processes: Leaf to Globe. pp. 7–19. Edited by J.R. Ebleringer and C.B. Field. Academic Press, San Diego, California.

Loveland, T.R., Merchant, J.W., Ohlen, D.O. and Brown, J.F. 1991. Development of a landcover characteristics database for the conterminous U.S. Photogrammetric Engineering and Remote Sensing 57: 1453–1463.

Meentemeyer, V. 1989. Geographical perspectives of space, time, and scale. Landscape Ecology 3: 163–173.

Milne, B.T. 1988. Measuring the fractal geometry of landscapes. Applied Mathematics and Computing 27: 67–79.

Nellis, M.D. and Briggs, J.M. 1989. The effect of spatial scale on Konza landscape classification using textural analysis. Landscape Ecology 2: 93–100.

O'Neill, R.V., DeAngelis, D.L., Waide, J.B. and Allen, T.F.H. 1986. A hierarchical concept of ecosystems. Princeton University Press, Princeton, New Jersey.

Openshaw, S. 1977. Optimal zoning systems for spatial interaction models. Environment and Planning A: 169–184.

Openshaw, S. 1984. The modifiable areal unit problem. CATMOG 38. GeoBooks, Norwich, England.

Openshaw, S. and Taylor, P. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *In:* Statistical Applications in the Spatial Sciences. pp. 127–144. Edited by N. Wrigley. Pion, London.

Openshaw, S. and Taylor, P.J. 1981. The modifiable areal unit problem. *In:* Quantitative Geography: A British View. pp. 60–69. Edited by N. Wrigley and R. Bennett. Routledge and Kegan Paul, London.

Openshaw, S., Charlton, M., Wymer, C. and Craft, A. 1987. Mark I geographical analysis machine for the automated analysis of point data sets. International Journal of Geographic Information Systems 1: 35–5 1.

Openshaw, S., Charlton, M. and Craft, A. 1988. Searching for Leukemia clusters using a geographical analysis machine. Papers of the Regional Science Association 64: 95–106.

Gehlke, C.E. and Biehl, K. 1934. Certain effects of grouping upon the size of the correlation coefficient in census tract material. Journal of the American Statistical Association Supplement 29: 169–170.

Pacala, S.W. and Silander Jr., J.A. 1985. Neighborhood models of plant population dynamics: I. Single–species models of annuals. American Naturalist 125: 385411.

Putnam, S.H. and Chung, S.-H. 1989. Effects of spatial system design on spatial interaction models. 1: The spatial system definition problem. Environment of Planning A 21: 2746.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V. and Hobbie, J.E. 1992. Aggregating fine-scale ecological knowledge to model coarser–scale attributes of ecosystems. Ecological Applications 2: 55–70.

Robinson, A.H. 1950. Ecological correlation and the behaviour of individuals. American Sociological Review 15: 351–357.

Rosswall, T., Woodmansee, R.G. and Risser, P.G., editors. 1988. Scales and Global Change. John Wiley & Sons, New York.

Roughgarden, J., Running, S.W. and Matson, P.A. 1991. What does remote sensing do for ecology? Ecology 7 2 4: 45–54

Taylor, P.J. and Johnston, R.J. 1979. Geography of Elections. Penguin, Harmondsworth.

Tobler, W.R. 1989. Frame independent spatial analysis. *In*: Accuracy of Spatial Databases, pp. 115–122. Edited by M.F. Goodchild and S. Gopal. Taylor and Francis, London.

Townshend, J.R.G. and Justice, C.O. 1990. The spatial variation of vegetation changes at very coarse scales. International Journal of Remote Sensing 11: 149–157.

Turner, M.G., Dale, V.H. and Gardner, R.H. 1989. Predicting across scales: theory development and testing. Landscape Ecology 3: 235–252.

Turner, S.J., O'Neill, R.V., Conley, W., Conley, M.R. and Humphries, H.C. 1981. Pattern and scale: Statistics for landscape ecology. *In*: Quantitative Methods in Landscape ecology: The Analysis and Interpretation of Landscape Heterogeneity. pp. 17–49. Edited by M.G. Turner and R.H. Gardner. Springer-Verlag, New York.

Turner, M.G., O'Neill, R.V., Gardner, R.H. and Milne, B.T. 1989. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3: 153–162.

Upton, G.J.G. and Fingleton, B. 1985. Spatial data Analysis by Example. Volume 1: Point Pattern and Quantitative Data. John Wiley & Sons, Chichester.

Urban, D.L., O'Neill, R.V. and Shugart, H.H. Landscape ecology. BioScience 37: 119–127.

Usher, M.B. 1975. Analysis of pattern in real and artificial plant populations. Journal of Ecology 63: 569–586.

Wiens, J.A. and Milne, B.T. 1989. Scaling of "landscapes" in landscape ecology, or, landscape ecology from a beetle's perspective. Landscape Ecology 3: 87–96.

Wiens, J.A., Stenseth, N.C., Van Home, B. and Ims, R.A. 1993. Ecological mechanisms and landscape ecology. **Oikos** 66: 369–380.

Williams, E. and Jelinski, D.E. 1995. On using the NOAA AVHRR "Experimental Calibrated Biweekly Global Vegetation Index". in press Photogrammetric Engineering and Remote Sensing.

Woodcock, C.E. and Strahler, A.H. 1987. The factor of scale in remote sensing. Remote Sensing of Environment 21: 311–332.

Wu, J. 1992. Detecting spatial patterns: the net–function interpolation. Coenoses 7: 137–143.

Wu, J. and Levin, S.A. 1994. A spatial patch dynamic modeling approach to pattern and process in an annual grassland. Ecological Monographs 64: 447–467.

Wu, J. and Loucks, O.L. 1995. From balance-of-nature to hierarchical path dynamics: a paradigm shift in ecology. Quarterly Review of Biology (in press).

Yule, ***G.*** and Kendall, M. 1950. An introduction to the theory of statistics. Charles Griffin and Company Limited, New York, New York, USA.