INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 2002
VOL. 16, NO. 1, 93–109

Taylor & Francis
Taylor & Francis Group

Research Article

Effects of spatial aggregation approaches on classified satellite imagery

HONG S. HE

School of Natural Resources, University of Missouri-Columbia, 203M ABNR Building, Columbia, MO 65211, USA; e-mail: Heh@missouri.edu

STEPHEN J. VENTURA

Environmental Studies and Soil Science, Land Information and Computer Graphics Facility, University of Wisconsin-Madison, 1525 Observatory Dr., Madison, WI 53706, USA

and DAVID J. MLADENOFF

Department of Forest Ecology & Management, University of Wisconsin-Madison, 1630 Linden Dr., Madison, WI 53706, USA

(*Received 10 June 2000; accepted 21 March 2001*)

Abstract. Spatial aggregations of raster data based on the majority rule have been typically used in landscape ecological studies. It is generally acknowledged that (1) dominant classes increase in abundance while minor classes decrease in abundance or even disappear through aggregation processes; and (2) spatial patterns also change with aggregations. In this paper, we examined an alternative, random rule-based aggregation and its effects on cover type abundance and landscape patterns, in comparison with the majority rule-based aggregation. We aggregated a classified TM imagery (about 1.5 million ha) from 30 m (4231×3717 pixels) incrementally to 990 m resolution (132 pixels $\times$ 116 pixels). Cover type proportion, mean patch size ratio, aggregation index (AI), and fractal dimension (FD) were used to assess the effects of aggregation. To compare landscapes under different resolutions, we assumed that the landscapes were least distorted if (1) the cover type proportions and mean patch size ratios among classes were maintained, and (2) all cover types responded in the same way for a given index as aggregation levels increased. For example, distortion is introduced by aggregation if some cover types increase their AI values with increasing aggregation levels while other cover types decrease. Our study indicated that the two spatial aggregation techniques led to different results in cover type proportions and they altered spatial pattern in opposite ways. The majority rule-based aggregations caused distortions of cover type proportions and spatial patterns. Generally, the majority rule-based aggregation filtered out minor classes and produced clumped landscapes. Such landscape representations are relatively easy for interpreting and, therefore, are suitable for land managers to conceptualize spatial patterns of a study region. By contrast, the random rule-based aggregations maintained cover type proportions accurately, but tended to make spatial patterns change toward disaggregation. Overall, the measurements of landscape indices used in this study indicated that

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810110075978

the random rule-based aggregation maintains spatial patterning better than the majority rule-based aggregation. Random rule-based aggregations are more suitable for studies in which the accuracy of spatially explicit information is of concern. They can be very useful in scaling data of fine resolution to coarse resolution, while retaining cover type proportions.

1. Introduction

Issues of biodiversity, wildlife conservation, and landscape change increasingly need to be examined at large spatial extents and multiple scales (resolutions). These often involve the use of satellite imagery covering large areas and the aggregation of the data from fine resolutions to coarse resolutions (Rastetter *et al.* 1992, Hess 1994, Bian and Butler 1999, Leibowitz and Hyman 1999). Aggregation is particularly important when data management and processing limitations constrain the use of fine resolution data, such as multiple full resolution Landsat scenes involving gigabytes of data.

For satellite data as well as other raster data, aggregations partition the input grid into blocks and determine the value for each block in the output grid. Such operations can be either numerical or categorical, depending on how the value for each output block is determined. Numerical aggregation involves generating new values by applying various mathematical methods to input data such as sum, mean, median, or a user-defined algorithm. For example, the mean is often used when aggregating imagery of NDVI to coarser resolutions (Townshend and Justice 1990, Jelinski and Wu 1996). Categorical aggregation generates output values by logical processing such as majority or predominant input category, or spatial processing such as assigning output values based on the input cell at the centre of the output grid cell. These are commonly used for classified satellite imagery or other thematic raster data.

Scientists have long noted data distortion or even data loss associated with aggregation or scaling (Gardner *et al.* 1982, Johnson and Howarth 1987, Woodcock and Strahler 1987, Hall *et al.* 1988, Wiens and Milne 1989, Levin 1992, Milne 1992). Efforts have been made to quantify such phenomena and to predict its effects across scales (Turner *et al.* 1989a, Baker 1993, Moody and Woodcock 1995, Collins and Woodcock 1999). While many studies have dealt with numerical aggregation such as the zoning or modifiable area unit problems (Fortheringham and Wong 1991, Jelinski and Wu 1996), some studies have used categorical aggregation (Moody and Woodcock 1995). The consensus from the studies of categorical aggregation is that (1) dominant classes become more dominant in terms of abundance or percentage cover, while minor classes become less common through aggregation processes; and (2) the rates of increase or decrease in abundance of various classes are strongly related to the spatial patterns, which also change with data aggregations. For example, Turner *et al.* (1989b) found that the abundance of minor land cover types decrease more slowly if they are clumped than if they are very dispersed as aggregation level increases. Moody and Woodcock (1995) found that classes with moderately sized patches tend to increase in abundance much more dramatically than classes with very large patches. While these studies suggest the possibility of predicting the distortion or loss of information based on initial proportions and spatial patterns of cover types (Turner *et al.* 1989a, Moody and Woodcock 1994, 1995), they have not examined alternative aggregation techniques.

Two common methods of categorical aggregation are based on either a majority or a random rule. The majority rule finds the majority value (the value that appears most often, but not necessarily >50%) for the specified cells of the input grid, and records that value in the corresponding aggregated cell on the output grid. Most studies examining categorical data aggregation were based the majority rule (Turner *et al.* 1989a, 1989b, Moody and Woodcock 1994, 1995, Benson and MacKenzie 1995, Mladenoff *et al.* 1997). Because of this rule, dominant classes are expected to become more dominant and minor classes decrease and in some cases disappear entirely from the landscape.

Random rule-based aggregations have not been evaluated as extensively as majority rule-based methods. The random rule assigns the value for the output cell by randomly picking a value among the specified cells of the input grid. Because the probability that a class is selected to represent an aggregated area is proportional to its presence in the input data, such aggregations are less likely to create data distortion or loss than those based on the majority rule, except under extreme conditions, such as when the starting map is aggregated to only a few pixels. Since cover type proportions of the random rule-based aggregation are scale independent, it is possible to examine directly the relationships between the levels of aggregation (scale) and resulting landscape patterns. Therefore, the objective of this paper is to compare the effects of majority and random rule-based aggregations. Specifically we intend to use a vegetation map based on a classified Landsat Thematic Mapper (TM) satellite image to examine the distortions introduced by data aggregation in terms of both cover type quantities and landscape patterns.

2. Methods

2.1. *Study area and the classified TM Landsat image*

Our study area is about 1.3 million ha and is located in northern Wisconsin, USA (44° N, 91° W) (figure 1). It comprises 62% forest, 10% lowland, 8% agriculture, 5% water, and 14% other land covers (Wolter *et al.* 1995). Forests in the study area are early successional (He *et al.* 1998) and fall in the transitional zone between boreal forest to the north and temperate forests to the south (Curtis 1959, Pastor and Mladenoff 1992). The forested areas have multiple ownerships, with about 40% state and national forest, 35% county forest, and 25% privately owned forest. Publicly owned forests tend to occur as large patches, while privately owned forests occur as small patches.

A full scene of Landsat TM satellite imagery of this region was classified at tree species level (Wolter *et al.* 1995). The sub-area used in the present study contains 4231 rows and 3717 columns of pixels with a map resolution of 30 m. For this study, we derived three major forest classes and a water class from the classified TM image. *Maple*, largely comprised of sugar maple (*Acer saccharum*), is the most common and clumped cover type, occurring on about 26% of the landscape. *Oak*, primarily comprised of northern red oak (*Quercus rubra*), is the least common and most dispersed cover type, occurring on about 4% of the landscape. *Mixed deciduous and coniferous* forest occurs as a moderately common (17%) and clumped cover type. It contains various combinations of conifer species including white pine (*Pinus strobus*), red pine (*P. resinosa*), and balsam fir (*Abies balsamea*), and deciduous species including yellow birch (*Betula alleghaniensis*), paper birch (*B. papyrifera*), and basswood (*Tilia americana*). *Water* occurs on about 5% of the landscape, similar in frequency to oak but it is the most clumped cover type.

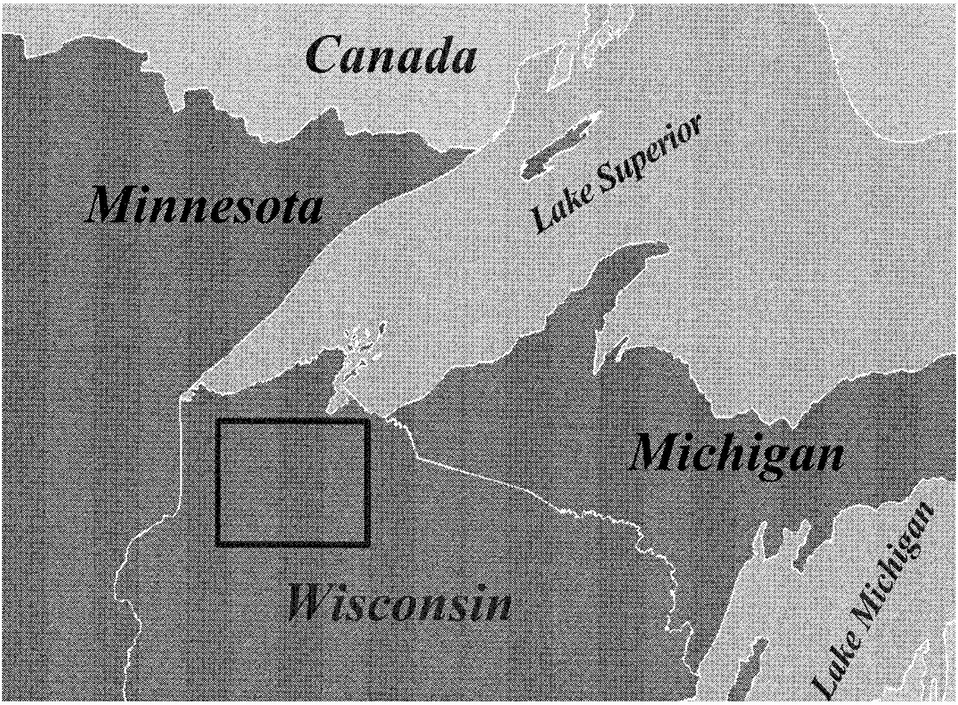

Figure 1. Study area, about 1.3 million ha, located in northern Wisconsin, USA (44° N, 91° W).

2.2. *Landscape metrics*

Landscape metrics commonly used to assess the effects of data aggregation include Shannon-Weaver diversity and evenness indices, dominance, contagion index, average patch size, and fractal dimension index (Turner *et al.* 1989b, Cullinan and Thomas 1992, Benson and MacKenzie 1995, Moody and Woodcock 1995, Wickham and Ritters 1995, Mladenoff *et al.* 1997). Diversity, evenness, and dominance indices are based on the proportion of each cover type on the landscape (Gustafson 1998, O'Neill *et al.* 1988), but inconsistent results in assessing spatial aggregation of different landscapes can be found by using these indices (Turner 1989b, Wickham and Ritters 1995, Moody and Woodcock 1995), because these indices do not incorporate spatial attributes, which are altered by data aggregations. The contagion index (*CI*) is a spatial index and was developed to measure the degree of clumpiness of overall landscape patterns (O'Neill *et al.* 1988). *CI* is sensitive to the number of cover types in the data set (Li and Reynolds 1993). It can be difficult to interpret and has been found to be inconsistent especially when linking the measurement to a specific cover type (Gardner and O'Neill 1991, Gustafson and Parker 1992, Wickham and Riitters 1995, Ritters *et al.* 1996, Schumaker 1996).

Average patch size is an intuitive index to measure aggregation and is especially suitable for categorical maps. Increasing levels of aggregation result in an increase in minimum patch sizes. Therefore, average patch sizes of all classes are expected to increase as well with the levels of aggregation. We calculated ratios of the average patch size among the four cover types.

Fractal dimension (*FD*) for a given cover type can be calculated using the area/

perimeter ratio relationship as two times the slope of the regression line found by regressing log (area) on log (perimeter) (Krummel *et al.* 1987). It has been found useful for assessing the effects of data aggregation (Wiens and Milne 1989, Cullinan and Thomas 1992, Jelinski and Wu 1996, Emerson *et al.* 1999). The complexity of patch shapes decreases with increasing levels of aggregation (pixel size), resulting in a decrease in *FD* value (Burrough 1981, McGarigal and Marks 1994).

He *et al.* (2000) developed a raster data based aggregation index (*AI*) similar to the class interspersion (*CL*) metric (Miller *et al.* 1997). *AI* for class *i* is derived as:

$$AI_i = e_{i,i}/max - e_{i,I}$$

where, $e_{i,i}$, the total edges shared by class *i* itself and $max - e_{i,i}$ is the largest possible number of shared edges for class *i*. Given a class *i* of area A_i, the maximum aggregation level is reached when *A* clumps into one patch that has the largest $e_{i,i}$ (e.g. it does not have to be a square). In raster data, if *n* is the side of largest integer square smaller than A_i, and $m = A_i - n^2$, then the largest number of shared edges for class *i*, $max - e_i$, will take one of the three forms:

$$max - e_{i,i} = 2n(n-1), \text{ when } m = 0, \text{ or}$$

$$max - e_{i,i} = 2n(n-1) + 2m - 1, \text{ when } m < n, \text{ or}$$

$$max - e_{i,i} = 2n(n-1) + 2m - 2, \text{ when } m > 0.$$

Alternatively, *AI* can be also calculated using the following single equation. Given a polygon containing *P* pixels let *X* be the edge length of the first square that fully contains area P. Then the denominator for *AI* is

$$AI = \text{int}\left(\frac{X^2 - P}{X}\right) - 2X + 2P \tag{1}$$

Note that the first term is done using integer math (so, for instance, it truncates 2.99 to 2). *AI* assumes that a class with the highest level of aggregation ($AI = 1$) is comprised of pixels sharing the most possible edges. A class completely disaggregated (pixels share no edges) has the lowest level of aggregation ($AI = 0$). *AI* is class specific, independent of landscape composition, and suitable for satellite data. It has been shown that *AI* can be used to compare classes from different landscapes, or the same class from a landscape at different resolutions (He *et al.* 2000). *AI* values of a given class larger than the corresponding lowest value would indicate a non-random pattern with some levels of aggregation (He *et al.* 2000).

In this study, we use *AI*, fractal dimension (*FD*), average patch size, as well as cover type proportions to assess the effects of data aggregation. To compare distortions on landscape patterns among different resolutions, we assume that landscape patterns are least distorted if (1) the proportions of cover types and the ratios of average patch sizes among classes are maintained somewhat constant, and (2) the fractal dimension and aggregation index measured for the four cover types respond in the same fashion (decrease or increase) to the levels of aggregations.

3. Aggregation scenarios

We aggregated the initial, classified TM satellite image based on either majority or random rule (hereafter referred to as RealM and RealR scenarios, respectively). The aggregations were sequential, from 30 to 990 m resolutions, a total of 32 iterations (30-m increment per iteration). Therefore, no pixel was divided by the aggregation procedure and data were not distorted owing to the grouping scheme. In addition, the range of spatial resolutions cover the span of most existing studies on data aggregations (Turner *et al.* 1989, Cullinan and Thomas 1992, Benson and MacKenzie 1995, Moody and Woodcock 1995, Wickham and Ritters 1995, Ritters *et al.* 1997) as well as the 250-m and 1000-m resolutions of Moderate Resolution Imaging Spectroradiometer (MODIS) (NASA's Earth Observing System 2000) and the Advanced Very High Resolution Radiometer (AVHRR), platforms used for broad-area ecological studies.

To minimize the impacts of spatial patterns and provide base-level comparisons to measurements from the scenario RealM and RealR, we randomized the pixels on the initial classified TM image but kept the proportion of each cover type unchanged. This created a random image with no spatial patterns. Corresponding to the RealM and RealR scenarios, we applied both majority and random rule-based aggregations on the randomized TM image. We did not find any comparative difference in any measurements for these two randomized map treatments. Therefore, only one of them was used, hereafter referred to as the Random scenario. For all scenarios, we measured cover type percentage area, *AI*, *FD*, and average patch size.

4. Results

4.1. *Percentage cover type*

For the RealM scenario, the abundance (percentage cover) of maple, the most common and contagious cover type, increased with the level of aggregation (increasing cell size). Such increases appeared to be more rapid for the initial aggregation iterations, while maple abundance increased by about 50% after six iterations when cell size reached 210 m. For the remaining 26 aggregation iterations, its abundance increased gradually by another 12% (figure 2(*a*)). Oak, the least common and most dispersed cover type, exhibited the opposite trend to maple. Its abundance decreased rapidly for the first few iterations of aggregation. When aggregated to 990-m resolution, only 22% of the original oak remained on the landscape (figure 2(*a*)). The abundance of water cover type maintained its level (5%) throughout the aggregation process (figure 2(*a*)). The abundance of the mixed deciduous and coniferous cover type lost to maple but was compensated from other minor classes. Therefore, it also remained relatively consistent during the aggregation process (figure 2(*a*)). These results are consistent with findings reported elsewhere where majority rule-based aggregation was examined (e.g. Turner *et al.* 1989, Moody and Woodcock 1995, He *et al.* 2000).

For the RealR scenario, abundance for all cover types remained unchanged, with average variations less than 1% (figure 2(*b*)). This suggests that no distortion and information loss in the amount of percentage cover be found when aggregating the classified TM satellite imagery from its finest resolution (30 m) to very coarse resolution (990 m). As expected, results of cover type percentage areas for the Random scenario remained constant across all scales examined and were identical to that of the RealR scenario (figure 2(*c*)).

4.2. *Aggregation index*

Before the aggregations, measurements of *AI* at the 30-m map resolution for the RealM and RealR scenarios indicated that maple had the highest *AI* value (0.54), while oak had the lowest *AI* (0.35). Water had a higher *AI* (0.47) than that of the mixed deciduous and coniferous class ($AI = 0.42$). For the Random scenario, the *AI*

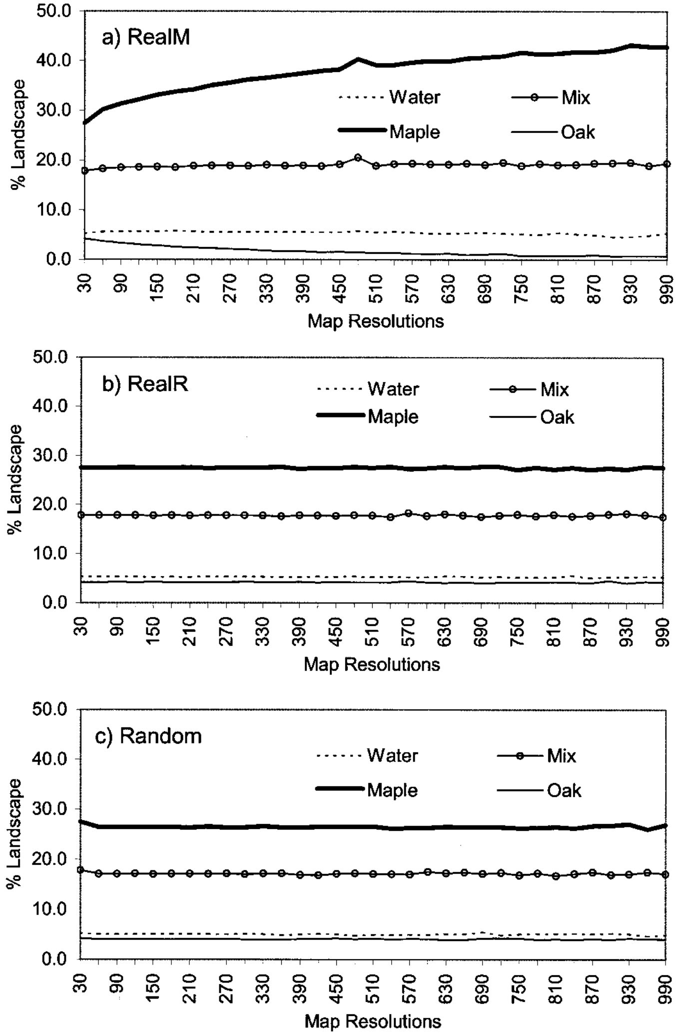

Figure 2. Cover type proportions of the classified TM satellite image aggregated from 30-m to 990-m map resolution under the scenarios of (*a*) the real landscape with the majority rule-based aggregation (RealM), (*b*) the real landscape with the random rule-based aggregation (RealR), and (*c*) the random landscape with random rule-based aggregation (Random).

of each cover type was significantly lower than the corresponding ones under the RealM and RealR scenarios. *AI*s of maple (0.26), mixed deciduous and coniferous (0.17), water (0.05), and oak (0.04) were the lowest possible and exactly corresponded to the proportions of cover types (percentage cover/100).

For the RealM scenario, *AI*s for the least common classes (water and oak) decreased and for the mid- and most common classes increased across all spatial scales examined (figure 3(*a*)). In other words, aggregation based on the majority rule made the dominant classes more aggregated and minor class more disaggregated (figure 4). For the Random scenario, *AI*s of all classes measured using the corresponding map resolutions remained constant (figure 3(*c*)). In these cases, scaling to a coarse

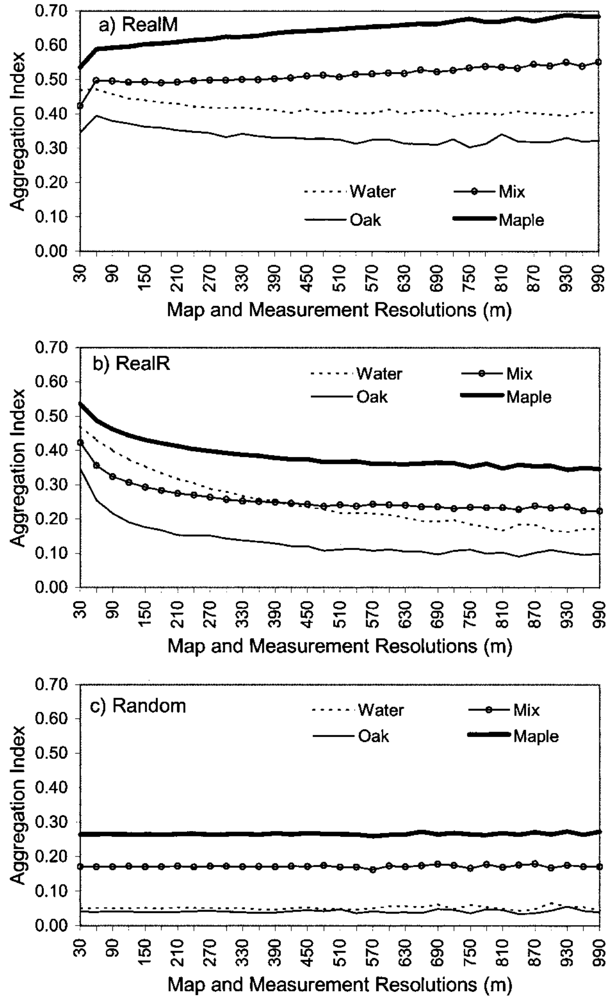

Figure 3. Aggregation Index (*AI*) measured using corresponding map resolutions for the classified TM satellite image aggregated from 30 to 990 m map resolution under the scenarios of (*a*) the real landscape with the majority rule-based aggregation (RealM), (*b*) the real landscape with the random rule-based aggregation (RealR), and (*c*) the random landscape with random rule-based aggregation (Random).

resolution, the resulting landscape becomes more aggregated when measured using the finer input map resolution. However, the resulting landscape is still random after the aggregation and, therefore, *AI* measured at the resulting map resolution remained unchanged and reflected the cover type proportions. This was seen across the spatial scales examined (figure 3(*c*)).

For the RealR scenario, the random rule-based aggregations randomly assigned values to output maps. This caused disaggregation of the existing cover type patches and resulted in less aggregated landscape patterns. Our results showed that *AI*s decreased for all classes, but more rapidly for the highly aggregated class (water) than for other classes (figure 3(*b*)). By 990-m resolution, *AI* decreased by 64%, 71%, 47%, and 36% for water, oak, mixed deciduous and coniferous, and maple, respect-

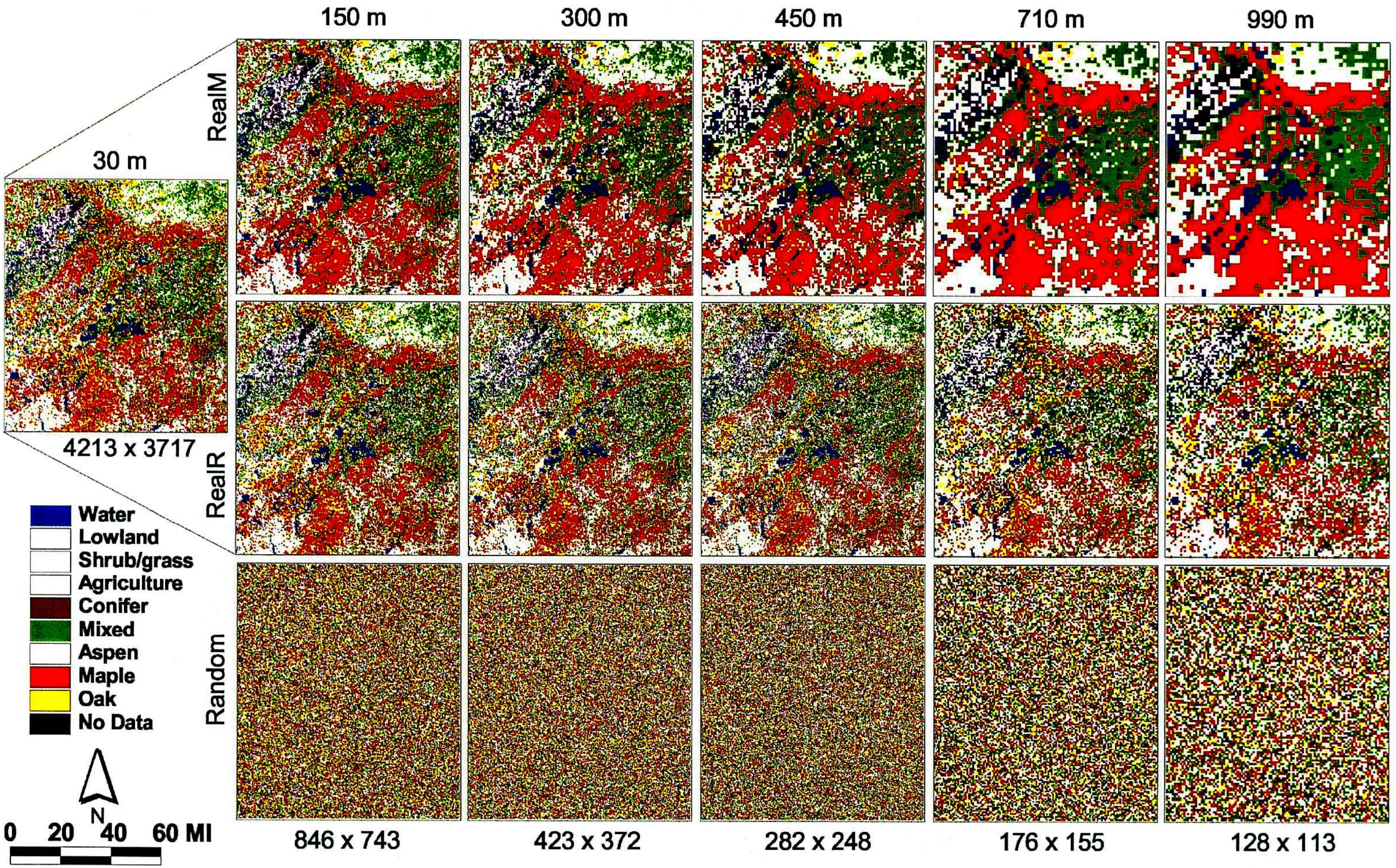

Figure 4. Maps of the classified TM satellite image aggregated from 30-m to 990-m map resolution under the majority and random rule-based scenarios. It is obvious that under the same resolution, maps under the random rule-based aggregation seem to have greater information than maps under the majority rule-based aggregation.

ively. The major decreases occurred during the initial iterations of aggregation. The rates of such decreases in *AI* became moderate as the level of aggregation increased.

4.3. *Fractal dimension*

Except for the most common class under the RealM scenario, *FD* measured across the range of resolutions decreased for all classes under both scenarios as the levels of aggregation increased. This is consistent with the results reported elsewhere (Milne 1988, O'Neill *et al.* 1991) suggesting that aggregations result in simplified spatial patterns (figures 5(*a–c*)). In the RealM scenario, the majority rule-based aggregation filtered out most small patches of minor classes and those left were in simple shapes (figure 5(*a*)). *FD* for maple remained relatively unchanged in contrast to decreases for all other classes for this scenario. This may be related to overall

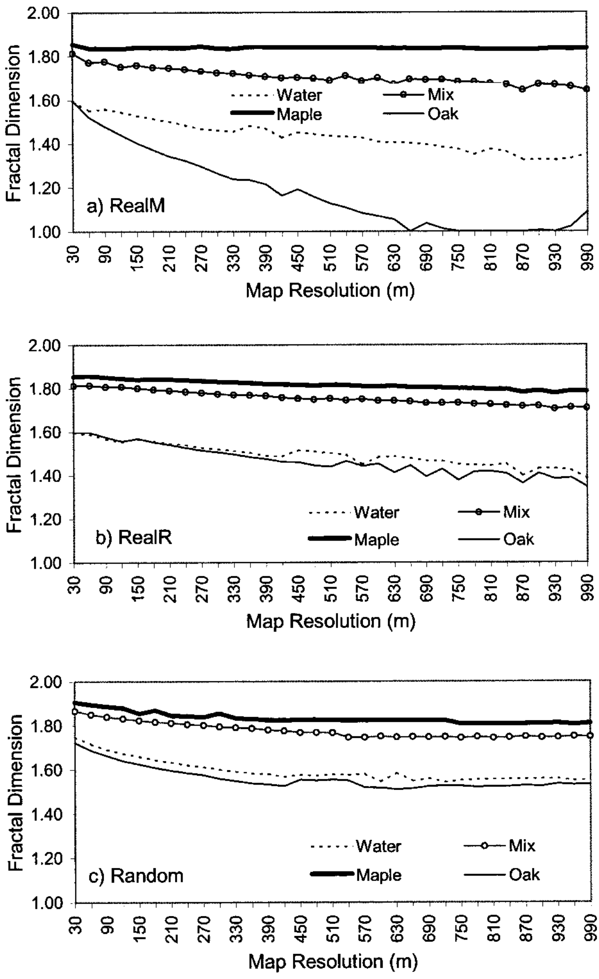

Figure 5. Fractal dimensions (*FD*) measured using corresponding map resolutions for the classified TM satellite image aggregated from 30 to 990 m map resolution under the scenarios of (*a*) the real landscape with the majority rule-based aggregation (RealM), (*b*) the real landscape with the random rule-based aggregation (RealR), and (*c*) the random landscape with random rule-based aggregation (Random).

landscape configurations that made *FD* of maple patches less variable than all other classes. For the RealR scenario, *FD*s of all four classes exhibited a near linear decreasing pattern with the increasing levels of aggregation (figure 5(*b*)). However, the slopes of such decreases are slower than those under the RealM scenario. This suggests that the RealR scenario creates less distortion in *FD* than the RealM scenario.

4.4. *Average patch size*

For all scenarios, average patch sizes of each class increased with the increasing levels of aggregation, as explained in the Methods section (figures 6(*a*, *b*)). Patch size variations reflected by the coefficients of variation (*cv*) decreased (figures 6(*c*, *d*)), indicating that aggregating to coarse resolutions decreased patch diversity. The average patch size coefficient of variation decreased more for the majority rule-based aggregation than for the random rule-based aggregation. For instance, in the RealM

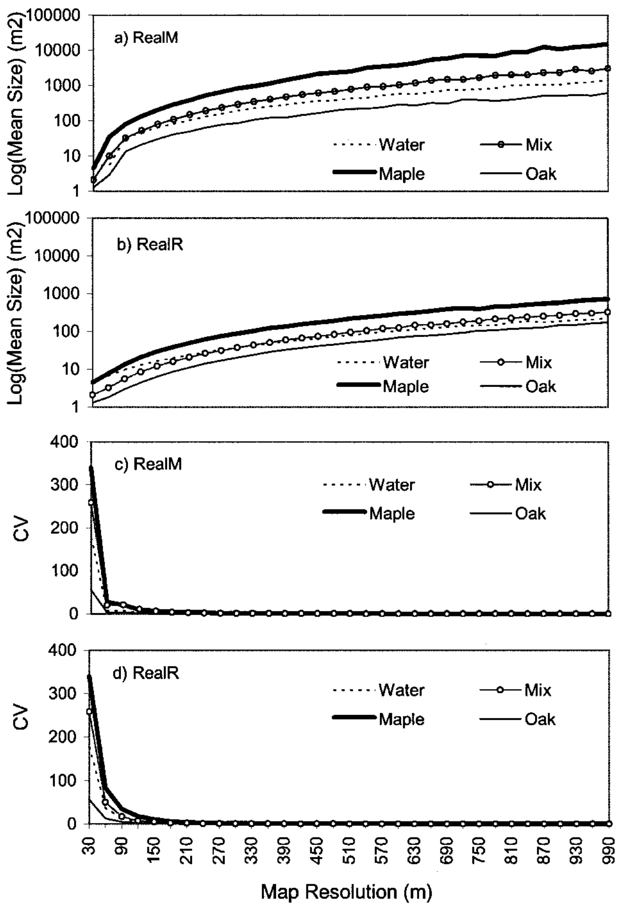

Figure 6. Mean patch size measured using corresponding map resolutions for the classified TM satellite image aggregated from 30-m to 990-m map resolution under the scenarios of (*a*) the real landscape with the majority rule-based aggregation (RealM), (*b*) the real landscape with the random rule-based aggregation (RealR), and coefficient of variation of mean patch size (*cv*) for (*c*) the real landscape with the majority rule-based aggregation (RealM), (*c*) the real landscape with the random rule-based aggregation (RealR).

scenario, the average patch size of water was 5.1 ha at 30-m resolution and its coefficient of variation was 176.70 (figure 6). After the first iteration of aggregation (at 60-m resolution), the average patch size increased to 5.8 ha, but the coefficient of variation decreased dramatically to 6.77. Such a decrease was much faster than the corresponding one ($cv = 38.03$ at 60-m resolution) under the RealR scenario. In general, the coefficient of variation diminished under increasing aggregation levels, with the RealR scenario maintaining higher *cv* values than the RealM scenario (figures 6(*c*, *d*)). To compare effects of aggregation on average patch sizes, we also calculated ratios of the average patch sizes among the four cover types (table 1). The RealM scenario altered the ratios of all cover types. Average patch size of maple increased more rapidly against all other classes as the levels of aggregation increased. At 30-m resolution, the ratio of water:maple is 1.14, mixed:maple 0.47, oak:maple 0.29; while at 990-m resolution, these ratios decreased to 0.10, 0.21, and 0.04, respectively. Water:maple and oak:maple decreased the most with the increasing levels of aggregation. It was obvious that the average patch sizes of less common classes such as oak and water increased more slowly against the medium and most common classes, with the ratios against the latter classes decreasing. The ratio of these two minor classes remained somewhat constant.

For the RealR scenario, there are only minor changes in average patch size ratios among mixed:maple, oak:maple, and therefore, oak:mixed. The only large change was the average patch size of water, which decreased more rapidly than all other classes. Water/maple ratio started at 1.14 at 30-m resolution and decreased to 0.32 at 990-m resolution (table 1). However, the water/maple ratio under the RealR scenario was still higher than that under the RealM scenario.

4.5. *Map fidelity through aggregation*

There is no doubt that information is lost during data aggregation. However, there is no single measurement of map fidelity, defined in this case as the correspondence of an aggregated image to the source image. The value of information preserved during data aggregation obviously depends upon how the derived maps are used. For example, if land cover proportion is the only concern, random rule-based aggregation does not change map fidelity.

A general evaluation of map fidelity can be done using correlation analysis between the original map and the aggregated maps (figure 7). The results indicate that map fidelity decreased substantially with increasing level of aggregation. Maps derived from the majority rule-based aggregation had lower fidelity than those under random rule-based aggregation, especially at the early steps of data aggregation. At the later steps, both methods show similar levels of decreasing map fidelity. At 990-m resolution, correlation coefficient (R) reaches about 0.33, indicating that about 67% of the original information is lost.

5. Discussion

Our study indicates that different spatial aggregation techniques across a broad range of spatial scales (30–990 m) lead to significantly different results in terms of cover type proportions and spatial patterns. Majority rule-based aggregations cause distortions of cover type percentage areas as has been previously found. The most common and contagious classes and the least common and dispersed classes are affected the most. By contrast, random rule-based aggregations do not produce such distortions, as proportions of various cover types remain relatively stable over

Table 1. Average patch sizes ratio and coefficients of variation (cv) of four cover types under the both majority and random rule-based aggregations.

Resolution (m)	RealM						RealR		
	Water/ maple	Mixed/ maple	Oak/ maple	Water/ mixed	Water/ oak	Oak/ mixed	Water/ maple	Mixed/ maple	Oak/ maple
30	1.14	0.47	0.29	2.42	3.87	0.62	1.14	0.47	0.29
60	0.17	0.29	0.08	0.57	1.99	0.29	0.87	0.42	0.23
90	0.41	0.39	0.16	1.04	2.48	0.42	0.73	0.41	0.22
120	0.35	0.38	0.15	0.91	2.27	0.40	0.63	0.41	0.21
150	0.32	0.39	0.15	0.81	2.14	0.38	0.57	0.41	0.22
180	0.29	0.38	0.14	0.75	2.03	0.37	0.51	0.41	0.22
210	0.27	0.39	0.13	0.70	2.10	0.33	0.49	0.42	0.22
240	0.25	0.36	0.12	0.71	2.10	0.34	0.45	0.42	0.23
270	0.24	0.35	0.12	0.68	2.07	0.33	0.44	0.41	0.23
300	0.23	0.34	0.10	0.68	2.28	0.30	0.43	0.43	0.23
330	0.24	0.36	0.11	0.66	2.15	0.31	0.42	0.43	0.23
360	0.22	0.35	0.11	0.63	2.11	0.30	0.40	0.41	0.23
390	0.20	0.33	0.09	0.61	2.31	0.26	0.41	0.45	0.24
420	0.19	0.32	0.08	0.60	2.28	0.26	0.39	0.43	0.23
450	0.17	0.29	0.08	0.60	2.24	0.27	0.40	0.44	0.24
480	0.17	0.29	0.08	0.57	1.99	0.29	0.39	0.44	0.24
510	0.17	0.31	0.09	0.55	2.00	0.28	0.37	0.43	0.23
540	0.14	0.28	0.07	0.51	2.08	0.24	0.37	0.44	0.23
570	0.15	0.27	0.07	0.57	2.20	0.26	0.38	0.46	0.23
600	0.16	0.27	0.08	0.57	2.04	0.28	0.35	0.42	0.24
630	0.13	0.27	0.06	0.49	2.13	0.23	0.36	0.46	0.24
660	0.12	0.25	0.06	0.48	2.10	0.23	0.35	0.43	0.23
690	0.13	0.25	0.05	0.51	2.42	0.21	0.33	0.41	0.22
720	0.11	0.21	0.06	0.51	1.87	0.27	0.34	0.43	0.22
750	0.11	0.24	0.06	0.48	2.05	0.24	0.37	0.48	0.27
780	0.12	0.28	0.05	0.44	2.35	0.19	0.32	0.46	0.23
810	0.12	0.23	0.05	0.52	2.60	0.20	0.37	0.47	0.25
840	0.12	0.22	0.05	0.52	2.36	0.22	0.36	0.47	0.24
870	0.09	0.19	0.04	0.46	2.05	0.22	0.32	0.46	0.23
900	0.10	0.22	0.05	0.46	2.14	0.22	0.33	0.46	0.25
930	0.10	0.23	0.04	0.42	2.19	0.19	0.32	0.47	0.23
960	0.10	0.20	0.04	0.49	2.44	0.20	0.30	0.44	0.24
990	0.10	0.21	0.04	0.47	2.35	0.20	0.32	0.45	0.24

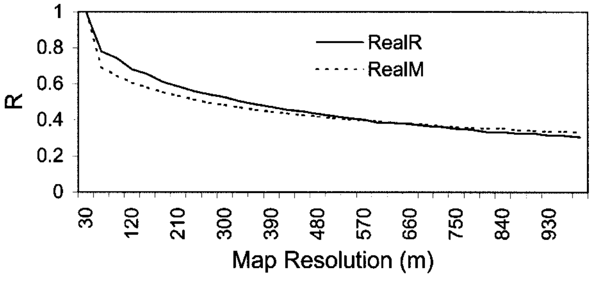

Figure 7. Correlation analysis of maps derived from the majority rule-based aggregation (RealM) has lower fidelity than those under random rule-based aggregation (RealR). at the early steps of data aggregation. At the later steps, both methods do not differ much.

30-m to 990-m resolutions and a large landscape (1.5 million ha, aggregated from 4231 × 3717 to 132 pixels × 116 pixels). This wide range of spatial resolutions is sufficient for most studies involving TM satellite data aggregations (Benson and MacKenzie 1995, Moody and Woodcock 1995, Wickham and Riitters 1995, Riitters *et al.* 1997). However, random rule-based aggregation cannot maintain constant cover type proportions indefinitely. Eventually, distortions would occur if maps were aggregated to a few pixels. We anticipate that distortion would not occur over even coarser resolution than 990 m if larger landscapes (> 1.5 million ha) are involved, consistent with conclusions reported elsewhere on the effects of spatial extents (Turner *et al.* 1989a).

Effects of aggregation on landscape patterns are complex, because landscape patterns cannot be represented by a single measure such as cover type proportion (O'Neill *et al.* 1988). A measurement of a landscape index is meaningful only when the measurement is taken under a resolution under which certain ecological processes (Wiens 1989, Schumaker 1996) can relate to the measurement results. The same map can be measured using different measurement resolutions and the results for a landscape index would differ (He *et al.* 2000). Since this study does not focus on the correspondence of measurement resolution (scales) and ecological processes, general comparisons of landscape metrics at variable map resolutions are employed.

In general, *AI*s measured with varied map resolutions for majority and random rule-based aggregations lead to the opposite results. The majority rule-based aggregations tend to make dominant classes clumped. Overall, *AI* does not dramatically decrease for the minor classes (figure 3(*a*)), because some tend to be disaggregated further by the majority rule, while others become more clumped as aggregation level increases. The random rule-based aggregations make all classes disaggregated. A decrease in *AI* indicates that a class is moving towards the state of randomness or no spatial pattern. Such a decrease suggests that random rule-based aggregation would eventually lead to a random landscape. However, the rates of such randomization appear to be significant only for the first a few iterations (e.g. 60–180 m), and gradually decreases with the increasing levels of aggregation (figure 3(*b*)). Since the lowest value of *AI* for a given class is equal to its cover type proportion measured for the random landscape, an *AI* value of a given class measured from the real landscape which is larger than the corresponding lowest possible value would indicate a non-random pattern with some degree of aggregation (He *et al.* 2000). In our study, at 990-m resolution, *AI*s for water, oak, mixed deciduous and coniferous derived under the RealR scenario are 272%, 153%, 31%, and 27% higher than of those of the random scenarios (figure 3(*c*)). This suggests that even at the highest aggregation level (990 m), considerable spatial patterning is maintained. These patterns are also apparent on the resulting maps (figure 4). Overall, random rule-based aggregations create less distortion in spatial pattern than the majority rule-based aggregations. This is also reflected in the results of fractal dimension and average patch size ratios. Average patch size ratios among classes are more constantly maintained and *AI* and *FD* of all four classes show more constant trends in random rule-based aggregations than those in majority rule-based aggregations (figures 3, 5 and 6).

It is difficult to disqualify one aggregation technique over the other. Choosing an aggregation method eventually depends upon the objectives. Majority rule-based aggregation has been widely used because many of the studies cited were aimed at simulating the type of land cover information that might be derived from coarser resolution source data (e.g. MODIS). It can filter out minor classes and produce clumped landscapes, and therefore, is suitable for regional ecosystem management in which plans such as ecological restoration and conservation are often conceptualized by managers based on the delineation of broad cover type distributions. Majority rule-based aggregation is also a built-in function of commonly used GIS software (ESRI 1998), while random rule-based aggregation can only be achieved through a customized program.

The random rule-based aggregation alters spatial patterns opposite to the majority rule-based aggregation, and may lead to disaggregated spatial patterns, which can make delineation of cover type distribution difficult. However, the random rule-based aggregation has been found superior to the majority rule-based aggregation for preserving several commonly used landscape metrics in this study. We have shown that the random rule-based aggregation maintains cover type proportions over a wide range of spatial scales, and maintains spatial information more precisely than the majority rule-based aggregation. At the same pixel size, a map from random rule-based aggregations seem to have finer resolution, since it portrays spatial information more accurately than the one from majority rule-based aggregations (figure 4). Therefore, the random rule-based aggregation can be more suitable for studies in which the accuracy of spatially explicit information is concerned. It can be very useful in scaling data of fine resolution to coarse resolution, while retaining cover type proportions.

Acknowledgments

We appreciate suggestions from, A-Xing Zhu, Ted Sickley, Harvey Miller, Janet Franklin, and the three anonymous reviewers. This work was funded in part through a cooperative agreement with the US Forest Service Northern Global Change Program and Mission Enhancement Fund from the University of Missouri-Columbia.

References

Baker, W. L., 1993, Spatially heterogeneous multi-scale response of landscapes to fire suppression. *Oikos*, **66**, 61–71.

Benson, B. J., and MacKenzie, M. D., 1995, Effects of sensor spatial resolution on landscape structure parameters. *Landscape Ecology*, **10**, 113–120.

Bian, L., and Butler, R., 1999, Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data. *Photogrammetric Engineering and Remote Sensing*, **65**, 73–84.

Burrough, P. A., 1981, Fractal dimensions of landscapes and other environmental data. *Nature*, **294**, 240–242.

Collins, J. E., and Woodcock, C. E., 1999, Geostatistical estimation of resolution-dependent variance in remotely sensed images. *Photogrammetric Engineering and Remote Sensing*, **65**, 41–50.

Cullinan, V. I., and Thomas, J. M., 1992, A comparison of quantitative methods for examining landscape pattern and scale. *Landscape Ecology*, **7**, 211–227.

Curtis, J. T., 1959, *The vegetation of Wisconsin: an ordination of plant communities.* (Madison, WI: University of Wisconsin Press).

Emerson, C. W., Lam, N. S.-N., Quattrochi, D. A., 1999, Multi-scale fractal analysis of image texture and pattern. *Photogrammetric Engineering and Remote Sensing,* **65**, 51–61.

ESRI, 1996, Arcview 3.0 (Redlands, CA: ESRI Inc.)

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

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

Gardner, R. H., and O'Neil, R. V., 1991, Pattern, process, and predictability: the use of neutral models for landscape analysis. In *Quantitative Methods in Landscape Ecology,* edited by M. G. Turner and R. H. Gardner (New York: Springer-Verlag), pp. 289–307.

Gustafson, E. J., 1998, Quantifying landscape patterns, What is the state of art? *Ecosystems*, **1**, 143–156.

Gustafson, E. J., and Parker, G. R., 1992, Relationship between landcover proportion and indices of landscape spatial pattern. *Landscape Ecology*, **7**, 101–110.

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.

He, H. S., Mladenoff, D. J., Radeloff, V. C., and Crow, T. R., 1998, Integration of GIS data and classified satellite imagery for regional forest assessment and landscape modeling. *Ecological Application*, **88**, 1072–1083.

He, H. S., DeZonia, B., and Mladenoff, D. J., 2000, An aggregation index to quantify spatial patterns of landscapes. *Landscape Ecology,* **15**, 591–601.

Hess, G., 1994, Pattern and error in landscape ecology: a commentary. *Landscape Ecology*, **9**, 3–5.

Jelinski, D. E., and Wu, J., 1996, The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, **11**, 129–140.

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

Krummel, J. R., Gardner, R. H., Sugihara, G., O'Neill, R. V., and Coleman, P. R., 1987, Landscape patterns in a disturbed environment. *Oikos,* **48**, 321–324.

Leibowitz, S. G., and Hyman, J. B., 1999, Use of scale invariance in evaluating judgement indicators. *Environmental Monitoring and Assessment*, **58**, 283–303.

Miller. J. N., Brooks, R. P., and Croonquist, M. J., 1997, Effects of landscape patterns on biotic communities. *Landscape Ecology*, **12**, 137–153.

Moody, A., and Woodcock, C. E., 1994, Scale-dependent errors in the estimation of land-cover proportions—implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing*, **60**, 585–594.

Moody, A., and Woodcock, C. E., 1995, The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing. *Landscape Ecology*, **6**, 363–379.

NASA's Earth Observing System 2000. MODIS, Moderate Resolution Imaging Spectroradiometer. http://modis.gsfc.nasa.gov/MODIS/ABOUT/

O'Neil, R. V., Krummel, J. R., Gardner, R. H., Sugihara, G., Jackson, B., Turner, M. G., Zygmunt, B., Christensen, S. W., Dale, V. H., and Graham, R. L., 1988, Indices of landscape pattern. *Landscape Ecology*, **1**, 153–162.

O'Neill, R. V., Turner, S. J., Cullinan, V. I., Coffin, D. P., Cook, T., Conley, W., Brunt, J., Thomas, J. M., Conley, M. R., and Gosz, J., 1991, Multiple landscape scales: an intersite comparison. *Landscape Ecology*, **5**, 137–144.

Pastor, J., and Mladenoff, D. J., 1992, The southern boreal-northern hardwood forest border. In *A systems analysis of the global boreal forest,* edited by R. L. Shugert and G. B. Bonan (Cambridge: Cambridge University Press), pp. 216–240.

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

Ritters, K. H., O'Neill, R. H., Wickham, J. D., and Jones, K. B., 1996, A note on contagion indices for landscape analysis. *Landscape Ecology*, **11**, 197–202.

Ritters, K. H., O'Neill, R. H., and Jones, K. B., 1997, Assessing habitat suitability at multiple scales: a landscape-level approach. *Biological Conservation*, **81**, 191–202.

Schumaker, N. H., 1996, Using landscape indices to predict habitat connectivity. *Ecology*, **77**, 1210–1225.

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., 1989a, Predicting across scales: theory development and testing. *Landscape Ecology,* **3**, 235–252.

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

Wickham, J. D., and Ritters, K. H., 1995, Sensitivity of landscape metrics to pixel size. *International Journal of Remote Sensing*, **16**, 3585–3594.

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

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

Wolter, P. T., Mladenoff, D. J., Host, G. E., and Crow, T. R., 1995, Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery. *Photogrammetric Engineering & Remote Sensing*, **61**, 1129–1143.

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