Effects of geographic information system vector–raster–vector data conversion on landscape indices[1]

Pete Bettinger, Gay A. Bradshaw, and George W. Weaver

Abstract: The effects of geographic information system (GIS) data conversion on several polygon-and landscape-level indices were evaluated by using a GIS vegetation coverage from eastern Oregon, U.S.A. A vector–raster–vector conversion process was used to examine changes in GIS data. This process is widely used for data input (digital scanning of vector maps) and somewhat less widely used for data conversion (output of GIS data to specific formats). Most measures were sensitive to the grid cell size used in the conversion process. At the polygon level, using the conversion process with grid cell sizes of 3.05, 6.10, and 10 m produced relatively small changes to the original polygons in terms of ln(polygon area), ln(polygon perimeter), and 1/(fractal dimension). When grid cell size increased to 20 and 30 m, however, polygons were significantly different ($p < 0.05$) according to these polygon-level indices. At the landscape level, the number of polygons, polygon size coefficient of variation (CV), and edge density increased, while mean polygon size and an interspersion and juxtaposition index (IJI) decreased. The youngest and oldest age-class polygons followed the trends of overall landscape only in terms of number of polygons, mean polygon size, CV, and IJI. One major side effect of the conversion process was that many small polygons were produced in and around narrow areas of the original polygons. An alleviation process (referred to as the dissolving process) was used to dissolve the boundaries between similarly attributed polygons. When we used the dissolving process, the rate of change for landscape-level indices slowed; although the number of polygons and CV still increased with larger grid cell sizes, the increase was less than when the dissolving process was not used. Mean polygon size, edge density, and fractal dimension decreased after use of the dissolving process. Trends for the youngest and oldest age-class polygons were similar to those for the total landscape, except that IJI was greater for these age-classes than for the total landscape.

Résumé : Les conséquences de la conversion de données d'un système d'information géographique (SIG) sur quelques indices polygonaux et paysagers ont été évaluées au moyen d'un SIG sur la végétation de l'Est de l'Orégon, aux États-Unis. Un processus de conversion vectoriel–matriciel–vectoriel a été employé afin d'examiner les changements dans les données des SIG. Ce processus est couramment utilisé pour la saisie de données (balayage de cartes vectorielles) et dans une certaine mesure pour la conversion de données (production de données de SIG en formats spécifiques). La plupart des valeurs mesurées étaient sensibles à la dimension des cellules matricielles utilisées lors de la conversion. Au niveau polygonal, l'utilisation de cellules mesurant 3,05, 6,10 et 10 m amenait relativement peu de modifications aux superficies, périmètres et dimensions fractales des polygones originaux. Si on portait la dimension des cellules à 20 et 30 m cependant, les polygones étaient affectés de façon significative ($p < 0{,}05$) selon les indices polygonaux. Au niveau paysager, on notait un accroissement du nombre de polygones, du coefficient de variation de la dimension des polygones (CV) et de la densité limitrophe alors que la dimension moyenne des polygones et l'indice de maillage et de juxtaposition (IMJ) diminuaient. Les polygones des classes d'âge inférieure et supérieure suivaient la tendance du paysage global uniquement en termes du nombre de polygones, de la dimension moyenne des polygones, de CV et IMJ. Un effet secondaire notable du processus de conversion a été que plusieurs petits polygones ont été créés dans des zones étroites à l'intérieur et autour des polygones originaux. On a utilisé un processus de fusion pour éliminer les frontières entre des polygones possédant des attributs similaires. Avec ce processus, le taux de changement des indices paysagers a diminué et bien que le nombre de polygones et le CV continuaient à augmenter avec les cellules plus grandes, l'accroissement était moindre que si le processus de fusion n'était

Received August 16, 1995. Accepted January 24, 1996.

P. Bettinger.[2] Department of Forest Resources, Oregon State University, Corvallis, OR 97331, U.S.A.
G.A. Bradshaw. USDA Forest Service, Pacific Northwest Research Station, Corvallis, OR 97331, U.S.A.
G.W. Weaver. Department of Forest Science, Oregon State University, Corvallis, OR 97331, U.S.A.

[1] Paper 3086, Forest Research Laboratory, Oregon State University, Corvallis.
[2] Author to whom all correspondence should be addressed.

pas utilisé. La dimension moyenne des polygones, la densité limitrophe et la dimension fractale diminuaient après le processus de fusion. Les tendances pour les polygones de classes d'âge inférieure et supérieure étaient les mêmes, sauf que la valeur de IMJ pour ces classes d'âge était supérieure à celle de l'ensemble du territoire.

[Traduit par la Rédaction]

Introduction

In contrast with the previous 80 years, forest resource planning on United States federal lands in the future will probably entail the explicit inclusion of spatial analysis techniques concerning natural resource phenomena. Geographic information systems (GIS), remote sensing technologies, and associated database developments will be used to meet this goal. U.S. national forest planning is moving to a position where spatial analysis is vital to successfully model the linkage between ecosystem health and land management activities. The Forest Ecosystem Management Assessment Team (Forest Ecosystem Management Assessment Team 1993) of the Pacific Northwest has emphasized the need for tracking the spatial relationships between many types of flora and fauna and land management activities and natural events, across several scales. Such complex analyses are facilitated by various computer-based methods, including GIS, remote sensing, and decision-support systems (Schuster et al. 1993). However, a limiting factor in the decision-making process is the ability to assimilate pertinent information in a timely and accurate manner (Loh and Rykiel 1992).

In this study, we focus on the problem of data conversion and its effects on polygon- and landscape-level indices. At some resolutions, data conversion may involve generalization, a process that smooths and reduces the number of vertices in a vector map. The representation of polygons and the landscape as a whole can thus change, which may affect such measurements as the amount of edge. The conversion process we use converts a vector map to a raster map and then converts the raster map back to a vector map. This process is a common form of data input for GIS, where analog vector maps are scanned (rasterized), then converted to vector GIS images. The process is also used for GIS data output to specific formats. For example, to convert a MOSS vector image to a SNAP II+ (Sessions and Sessions 1993) data format, a MOSS vector image is rasterized by using LTPlus (USDA Forest Service 1992) and is subsequently vectorized and formatted for use in SNAP II+. It is this data output process that we examine in this study. When different grid cell sizes are used in the data conversion process, GIS data sets become generalized to various extents.

We use five raster cell sizes in the conversion process and examine the effects of conversion on the number of polygons, mean polygon size, polygon size coefficient of variation, fractal dimension, edge density, and an interspersion–juxtaposition index. Our goal is to determine whether these indices are sensitive to changes in the grid cell size used. Because this study focuses on the effects of data conversion on polygon- and landscape-level indices, we provide a brief discussion of error propagation in GIS data conversion, as well as the six polygon- and landscape-level indices of interest, prior to presenting the methods, results, and discussion of our research.

Optimally, data in information systems should be used to support decisions at the level the data were collected (e.g., stand-level data supporting tactical plans) (Burrough 1989). In practice, however, data are frequently aggregated at much broader scales to accommodate regional analyses. The level at which data are collected is a compromise forced by the constraints of time and money and molded to the proximal goals for which the information was collected (Burrough 1989). There are conceptual obstacles to surmount regarding scale. A determination should be made of those processes that remain constant and those that change across scales; there may be a wide range of spatial scales for landscape parameters, in which biophysical processes may vary. For example, Benson and Mackenzie (1995) report that several landscape parameters are sensitive to grain sizes, which increase from 20 to 1100 m in remotely sensed data.

Error propagation in GIS data conversion

The propagation of error caused by standard map operations has been thoroughly investigated (Haining and Arbia 1993). Walsh et al. (1987) divide errors into two types: (*i*) inherent and (*ii*) operational. Inherent error is present in the source documents (e.g., field measurement error) and is derived from source map projection, map construction techniques, and symbolization of data. Operational error is produced through the data capture and manipulation procedures in GIS; it is derived from data entry, manipulation, extraction, and comparison processes. This study is primarily concerned with operational error in data conversion from vector to raster to vector data structures; for example, data input (i.e., digital scanning) or data output, depending on the GIS system being used.

The vector to raster data conversion process includes importing vector data into the computer program for processing, applying the appropriate scales and map transformations, and then gridding the vector data (Figs. 1*a*–1*c*). Much of the research on rasterizing error evaluates the percentage of expected variance for each grid cell, or the thematic error (Wehde 1982; Clarke 1985; Veregin 1989; Valenzuela and Baumgardner 1990; Bregt et al. 1991; Carver and Brunsden 1994). Carver and Brunsden (1994) found that grid cell size has more effect than grid orientation or the rasterization process on rasterizing error. Errors are the largest when grid cell size is large in relation to polygon size and when the boundary of the vector polygon is complex. As grid cells become smaller relative to the conditions they represent, measurements based on their centers become more precise (Tomlin 1990). Generally, the smaller the grid cell size, the smaller the operational error. Another argument against using large grid cells is that spatial variation on an order of magnitude below the grid cell size is

Fig. 1. The vector to raster process: (*a*) vectors are imported into the computer program; (*b*) vectors are "gridded," where grid cells that touch the vectors are retained; and (*c*) the resulting raster image is saved. The raster to vector (vectorization) process: (*d*) the raster image is thinned; (*e*) lines are formed by connecting the centers of each grid cell; and (*f*) the resulting vector image is saved.

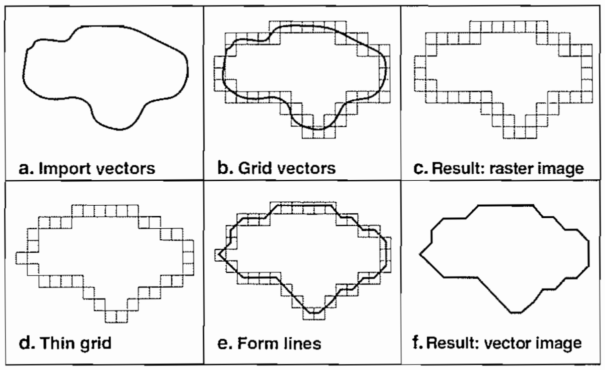
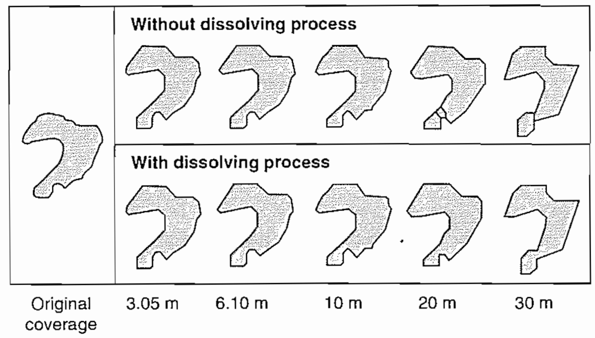


Fig. 2. Representation of original coverage at five grid sizes, without and with a dissolving process applied.



lost, which at times leads to highly generalized data layers; distortions in the shape and location of polygon boundaries become quite significant (Clarke 1985; Veregin 1989). However, using very small grid cells represents a potentially wasteful use of computer storage capacity.

The demand for conversion of raster data to vector format (vectorizing) has increased as raster data systems (e.g., remote sensing) have become more widely used (Clarke 1990). Government and nongovernment organizations are spending large sums of money acquiring and classifying remotely sensed data to produce land use, land cover, and other thematic GIS databases for purposes of conservation, inventory, and land use decisions (Hess 1994). Vectorizing processes are generally very CPU intensive. Topological and other errors may appear in the resulting vector data set. The process begins with line thinning, where the grid is thinned to a set of grid cells that is one cell wide (Fig. 1*d*). Vertices, usually located in the center of the remaining grid cells, are then connected to form lines, and the topology of the resulting vector data set is then reconstructed (Figs. 1*e*–1*f*) (Clarke 1990). When a raster grid is too coarse relative to map features, "bridging" may connect previously separated lines, dividing narrow portions of a polygon into several small polygons (Fig. 2). Typically, bridging problems occur where map features are close together relative to the grid cell size used. A dissolving process (eliminating the boundary between similar polygons) is one possible way to alleviate the cartographic problems encountered in the conversion process.

Polygon-level and landscape-level indices

The size, shape, distribution, and complexity of patches are functions of both natural disturbance processes and land management activities. Processes such as the spread of disturbances and species, the distribution, movement, and persistence of species, and the patterns of production and redistribution of matter and energy can be understood by examining landscape patterns. These patterns have been described in terms of dominance, contagion, complexity, edges, and the size, shape, and distribution of patches. The significance of polygon- and landscape-level indices varies with respect to ecosystem health and the effects of past or future land management activities. Although many indices could have been considered, we have chosen six indices to illustrate the effects of conversion. Our results are specific to a case study area, a portion of the Wallowa-Whitman National Forest in eastern Oregon, U.S.A. However, we feel the trends could be important to other regions of the world, and we point out the dangers in assuming that spatial resolution of data has no effect on the results of data conversion.

Patch size is a major variable that affects total biomass, production, and nutrient storage per unit area, in addition to species composition and diversity (Forman and Godron 1986). Small patches may contribute to local and regional extinction of sensitive species that are adapted to interior conditions (Dunn et al. 1991); they may also harbor exotic or weedy species because of their greater edge to interior ratio (Dunn and Loehle 1988). Further, patch size influences aesthetics and the microclimatic conditions within the patch (Bradshaw 1992) and may influence bird species diversity (McIntyre 1995). From a planning perspective, patch size influences operational practices such as the use of fire, regeneration establishment, stand development, and economics; however, no single patch size is ideal for all silvicultural requirements (Bradshaw 1992). Thus the significance of patch size varies with land management objectives and species habitat requirements. Polygon size coefficient of variation (CV) is a relative measure of dispersion expressed as a percentage of the units of the particular data; it is useful for comparing the variability of two or more sets of data (Berenson and Levine 1983) and may be indicative of the disturbance regime that structured the landscape.

Patch shape is an important factor in describing landscape; its significance is related to edge effect (Forman and Godron 1986). The way shape differs with scale can be characterized by fractal dimension (*D*). Fractal theory is based on two tenets (that phenomena exhibit dimensionality on a continuum and that fractal features are self-similar across a wide range of scales (Cartensen 1989)). Fractals are used for examining similarities of geographic patterns at different scales and quantifying line shape and surface roughness (Cartensen 1989). A constant D across scales indicates statistical self-similarity, and a change in D at different scales may indicate the dominance of different processes or constraints operating at different scales (Wiens 1989). D is not expected to be constant in reality, except in small areas over limited ranges of scale (Goodchild 1980). D allows ecologists to view the landscape at multiple scales and thereby achieve predictability in the face of ambiguity (Milne 1991). The area to perimeter relationship of D can provide insight into the nature of fragmentation in polygon coverages (Krummel et al. 1987), but scale must be appropriately defined if specific spatial patterns are to be evaluated (Ripple et al. 1991).

Changes in edge may have important implications for species persistence. Hunter (1990) proposes that woodlot shape may affect bird species richness. An increase in edge may lead to an increase in the diversity and abundance of wildlife species, in general, across a landscape (Hunter 1990). Increases in edge, however, may not be beneficial for certain individual species, such as some forest-interior bird populations. Edge density is the sum of the lengths (m) of all edge segments for each age-class divided by landscape area (m^2) (McGarigal and Marks 1993). Landscape boundary segments and background edge segments are not used in the calculations in this study.

$$\text{edge density (m/ha)} = \left(\frac{\sum \text{edge}_{\text{age-class } i}}{\text{total landscape area}}\right) \times 10\,000$$

where i represents a predefined age-class group. Edge density has been used to evaluate disturbance rates (Franklin and Forman 1987), tree growth (Hansen et al. 1993), and predation (Gates and Gysel 1978).

The interspersion and juxtaposition index (IJI) measures the distribution of adjacencies among patch types (McGarigal and Marks 1993). In our example, it is the distribution of adjacencies among age-classes. IJI ranges from 0 to 100, where the lower range represents a situation of clumped age-classes (all age-classes are not equally adjacent to all other age-classes), and the upper range represents maximum interspersion and juxtaposition.

$$\text{IJI } (\%) = \frac{-\Sigma_{i=1}\Sigma_{k=i+1}\left(\frac{e_{ik}}{E} \times \ln\frac{e_{ik}}{E}\right)}{\ln\left(\frac{1}{2}[m(m-1)]\right)} \times 100$$

where

i and k are the number of patches of types i and k

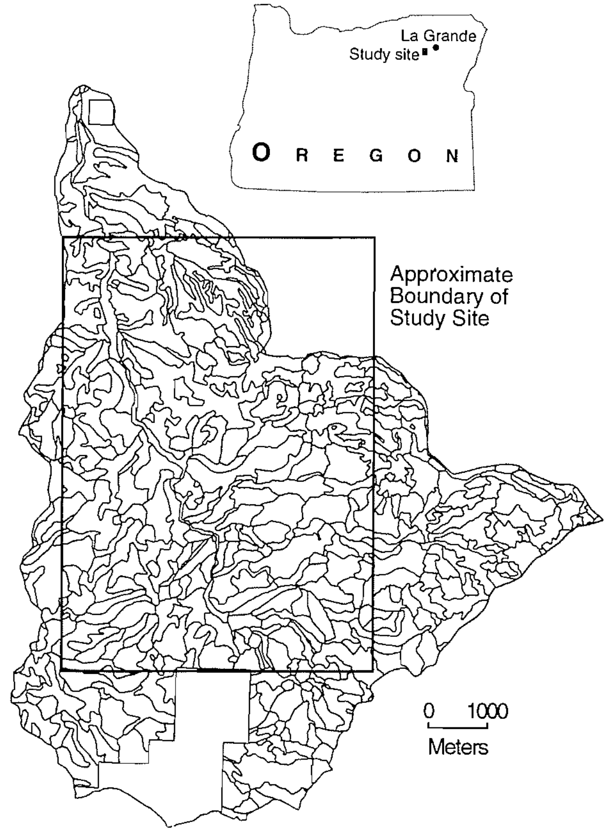
e_{ik} is the total length (m) of edge in landscape between patch types i and k

E is the total length (m) of edge of landscape

m is the number of patch types present in landscape

Although IJI is an important measure of adjacency within the landscape, the evaluation of the index from a management perspective obviously varies. IJI is similar to an alternative metric, the contagion index, proposed by Li and Reynolds (1994).

Fig. 3. Location of study area.



Study area

The Wallowa-Whitman National Forest in eastern Oregon is a good example to illustrate the implications of GIS data generalization on natural resource use. We focused on one particular area, the Upper Grande Ronde River Basin located primarily in the La Grande Ranger District, directly west of La Grande (Fig. 3). The basin is composed of a mixture of forested areas (with ponderosa pine (*Pinus ponderosa* Dougl. ex Laws.), Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), and other tree species) and rangeland. Range vegetation varies from riparian meadow bottoms to high alpine meadows with short growing seasons. The headwaters of the basin are composed of rugged mountains. Precipitation ranges from 51 to 97 cm per year; runoff is primarily snowmelt, with peak flows occurring in the spring. A study area of approximately 3.75′ by 3.75′ (one-fourth of a 7.5′ topographic map) was chosen for the case study.

Methods

The study area comprised 257 vegetation polygons totalling approximately 3162 ha. The GIS coverage of the study area was exported as a vector coverage from MOSS into LTPlus (USDA Forest Service 1992), where the map was rasterized at five different grid cell sizes. The resulting maps were then vectorized in LTPlus and "snapped" back to vertices and lines of the original coverage, to a maximum distance of half the grid cell size. This is a common process designed as an attempt to minimize deviations in vector lines from the original map, yet it is not entirely perfect because the new vertices may be located at a distance from the original vertices of more than one-half the grid cell size. Thus the differences between the original map and the converted map are a function of both the rasterizing and vectorizing processes. After the conversions, the coverages were exported to a digital exchange file format. These coverages were imported into ArcInfo (Environmental Systems Research Institute 1994) so that the effects of data conversion on landscape indices could be evaluated with FRAGSTATS (McGarigal and Marks 1993). Our objective was to simulate the data conversion process from MOSS to SNAP II+ (Sessions and Sessions 1993) data format. SNAP II+ is a tactical planning tool for forest management that can schedule land management activities subject to both spatial and nonspatial constraints.

Five square grid cell sizes were used in the data conversion process: 3.05 m (10 ft), 6.10 m (20 ft), 10 m (32.8 ft), 20 m (65.6 ft), and 30 m (98.4 ft). Because the location of the La Grande Ranger District's original 6.10-m grid was not documented during the data input process, 3.05 m was used as the minimum cell size. The 10-, 20-, and 30-m grid cell sizes were chosen to simulate satellite imagery resolution (i.e., SPOT and Thematic Mapper).

Two duplicate vector coverages were created in the data conversion process for each of the five grid cell sizes; this procedure gave 10 coverages for comparison with the original (preconversion) coverage. All 10 GIS coverages encompassed the same areal extent. One set of the five resolutions was subjected to an alleviation process ("dissolving") in order to eliminate boundaries between similarly attributed polygons caused by "bridging" across narrow areas of the polygons (Fig. 2). Similarly attributed polygons are defined as those with exact matches of codes for dominant tree species, age-class, and trees per acre. We present results for the entire landscape and for the youngest (0–10 years) and oldest (140+ years) age-classes; together these two age-classes represent more than 51% of the case study area. Age-classes were based on the age of the dominant overstory vegetation in each polygon, by 10-year intervals.

To assess the differences among the 10 converted coverages and the original coverage, we performed two analyses, one at the polygon level and one at the landscape level. At the polygon level, we calculated basic statistics concerning the size, perimeter, and D of the vegetation polygons. We evaluated the distribution of polygon size, perimeter, and D for normality and subsequently determined that transformations of each were required prior to performance of an analysis of variance (ANOVA) on the polygon-level data. Transformations on the polygon-level data were necessary to stabilize the variances of the data within each coverage, so that the variances between the coverages could be more effectively evaluated. Polygon size and perimeter were transformed by a natural logarithm transformation, and D was transformed by an inverse transformation. An ANOVA was used to determine whether there were significant differences among the two sets of GIS coverages (before dissolving and after dissolving): if the ANOVA signalled a significant difference within one set of GIS coverages (e.g., among the coverages before dissolving), pairwise t-tests were utilized to determine whether each combination of two coverages was significantly different.

Landscape-level indices were calculated with an ArcInfo (Environmental Systems Research Institute 1994) mediated software package, FRAGSTATS (McGarigal and Marks 1993). The landscape-level indices chosen were number of polygons, mean polygon size, CV, D, edge density, and IJI. Because only one measure was reported for each landscape-level index, a determination of the statistical differences among coverages was not possible. Therefore, we report the results and discuss only the possible trends in the data at the landscape level.

The use of adjacent polygons in the ANOVA could possibly violate the assumption of independence of data, because changes in a metric of one polygon may affect the same metric of an adjacent polygon. An analysis of the correlation of changes in a "seed" polygon and associated changes in adjacent polygons was used to evaluate whether adjacent polygons could be considered independent samples. For example, a change in the metric of polygon α between the original and the 3.05-m coverages was used to examine the correlation with changes in adjacent polygons β and δ or

$$(\alpha_{\mathrm{orig}} - \alpha_{3.05\ \mathrm{m}}) \text{ with } (\beta_{\mathrm{orig}} - \beta_{3.05\ \mathrm{m}}) \text{ and with } (\delta_{\mathrm{orig}} - \delta_{3.05\ \mathrm{m}})$$

The hypothesis being tested is that the correlation between changes in adjacent polygons is not significant. If changes are negatively correlated (e.g., if the metric for one polygon increases while it decreases for an adjacent polygon), then adjacent polygons are not independent with regard to the metric under consideration.

The preceding analysis does not examine how individual polygons change as a result of data conversion, because polygons can be split into several smaller polygons, each treated as an individual sample in the ANOVA. We therefore selected a random sample of polygons, summed their area and perimeter, and calculated D for each GIS coverage. We then developed regression equations in an attempt to predict changes in polygon area, perimeter, and D for polygons originally less than 10 ha, 10–20 ha, and greater than 20 ha. A stepwise procedure was used, with the dependent variables consisting of the differences between the original and the converted polygon metrics. Predictor variables included ln(original polygon area), ln(original polygon perimeter), (1/original D), and indicator variables (0–1) representing each grid cell size. The probability of entering values and of removing values was set at 0.05. We also attempted to develop exponential, non-linear regression equations but found none of them to be better predictors of changes in polygon metrics than the linear models.

Results

Changes in area, perimeter, and D of adjacent polygons showed no significant negative correlation when metrics for the original and all 10 of the converted coverages were examined; for example, the degree of negative association in pairs of variables (i.e., $(\alpha_{\mathrm{orig}} - \alpha_{30\ \mathrm{m}})$ with $(\beta_{\mathrm{orig}} - \beta_{30\ \mathrm{m}})$) is not significant. The only significant positive correlation ($p < 0.05$) occurred when a 30-m grid cell size with no alleviation process was used during data conversion. In this case, adjacent polygons increased or decreased in perimeter together. What this suggests is that during data conversion it is not the size or shape of one polygon that affects the size or shape of an adjacent polygon. Rather, it is probably the size of the grid cell used in the data conversion, the location of an original vector line just prior to rasterization, or the location of a vector line after vectorization that affected polygon metrics in this case study. Thus the utilization of adjacent polygons in the ANOVA probably does not violate the assumption of independence of the sample data.

Polygon- and landscape-level results are reported for coverages in which the dissolving process was not used and then for coverages in which it was applied. Results show that most measures are sensitive to the grid cell size used in the generalization process; however, after dissolving was applied, the measures were not as sensitive to grid cell size.

Table 1. ANOVA table of six GIS coverages.

Variable	Mean	Source of variation	df	Sum of squares	Mean squares	F-value	p-value
Without dissolving							
Ln(polygon area)	1.392 ha	Among groups	5	647.800	129.560	56.93	0.0001
		Within groups	1942	4419.752	2.276		
		Total	1947	5067.552	2.603		
Ln(polygon perimeter)	6.964 m	Among groups	5	277.841	55.574	64.46	0.0001
		Within groups	1942	1648.807	0.849		
		Total	1947	1926.648	0.990		
Inverse (fractal dimension) ($1/D^a$)	0.760	Among groups	5	0.008	0.001 5	2.96	0.0115
		Within groups	1942	1.015	0.000 5		
		Total	1947	1.023	0.000 5		
With dissolving							
Ln(polygon area)	1.948 ha	Among groups	5	9.250	1.850	1.37	0.2330
		Within groups	1582	2137.575	1.351		
		Total	1587	2146.825	1.353		
Ln(polygon perimeter)	7.318 m	Among groups	5	5.583	1.117	2.05	0.0687
		Within groups	1582	860.417	0.544		
		Total	1587	866.000	0.546		
Inverse (fractal dimension) ($1/D^a$)	0.762	Among groups	5	0.003	0.000 56	1.16	0.3254
		Within groups	1582	0.763	0.000 48		
		Total	1587	0.766	0.000 48		

[a]Fractal dimension (D) = {2 × logarithm(perimeter)}/logarithm(area).

Polygon-level indices without use of the dissolving process

Without the dissolving process, the ANOVA for polygon-level data showed significant differences in ln(polygon area), ln(polygon perimeter), and $1/D$ between the original coverage and the five generalized coverages (Table 1). In evaluations of the paired t-tests for each combination of two coverages, no significant difference ($p > 0.05$) between the original coverage and the 3.05-, 6.10-, and 10-m coverages was found for any of the variables, but the 20- and 30-m coverages were significantly different ($p < 0.05$) from all other coverages in terms of ln(polygon area) and ln(polygon perimeter). For $1/D$, the 20-m coverage was significantly different ($p < 0.05$) from all others. Several polygons, such as the one shown in Fig. 2, comprised more pieces after the 20-m conversion than after either the 10- or 30-m conversions. These smaller polygons had a high D value, thus raising the average D for the coverage enough to make the 20-m coverage significantly different from all the others.

Landscape-level indices without use of the dissolving process

There was a dramatic increase in the number of polygons (and hence polygon density) for the total landscape when grid cells larger than 10 m were used (Table 2). The 20- and 30-m coverages became severely fragmented by the data conversion process, so that the total number of polygons after data conversion ranged from 272 (3.05-, 6.10-, and 10-m coverages) to 495 (30-m coverage). The mean size of the polygons decreased correspondingly, as did IJI; CV and edge density increased. D stayed relatively constant, except when the 20-m grid cell size was used. When polygons in only the youngest age-class (0–10 years old) were examined, the number of polygons, CV, and D increased with the larger cell sizes, but mean polygon size and IJI decreased (Table 2). Edge density did not show a definite trend. In the oldest age-class (140+ years old), the number of polygons and CV also increased beyond the 10-m grid cell size, while mean polygon size and edge density decreased. Here, D and IJI decreased substantially when the 30-m grid cell size was used.

Polygon-level indices after use of the dissolving process

The dissolving process reduced the large number of polygons found at large grid cell sizes with the nondissolved results and increased mean polygon size; statistical techniques no longer showed significant differences between the original coverage and the five converted coverages (Table 1). ANOVA results showed no significant differences in ln(polygon area), ln(polygon perimeter), or $1/D$ among the original coverage and the five converted coverages.

Landscape-level indices after use of the dissolving process

Even with the dissolving process, the number of polygons for the total landscape increased with larger cell sizes

Table 2. Landscape-level indices based on GIS vegetation coverages with no dissolving process applied.

Index[a]	Original coverage	Grid cell size used in data conversion				
		3.05 m	6.10 m	10 m	20 m	30 m
Total landscape						
Number of polygons	257	272	272	272	379	495
Mean polygon size (ha)	12.301	11.624	11.630	11.630	8.342	6.386
Polygon CV (%)	105.629	110.371	110.332	110.446	142.287	163.376
D^b	1.315	1.317	1.316	1.316	1.325	1.315
Edge density (m/ha)	74.545	74.335	74.319	74.441	74.852	74.949
IJI	69.034	68.940	68.935	68.917	68.414	66.812
Youngest age-class (≤10 years old)						
Number of polygons	42	48	48	48	61	81
Mean polygon size (ha)	10.181	8.909	8.915	8.910	7.007	5.257
Polygon CV (%)	83.289	91.158	91.257	91.151	114.513	136.040
D^b	1.306	1.308	1.308	1.308	1.317	1.312
Edge density (m/ha)	17.745	17.853	17.841	17.848	17.933	17.843
IJI	78.752	78.734	78.714	78.664	78.289	76.700
Oldest age-class (≥140 years old)						
Number of polygons	90	92	92	92	117	147
Mean polygon size (ha)	11.711	11.464	11.480	11.471	8.998	7.180
Polygon CV (%)	81.614	84.007	84.070	84.011	106.390	125.175
D^b	1.319	1.322	1.319	1.318	1.320	1.311
Edge density (m/ha)	39.020	38.881	38.884	38.956	38.737	38.321
IJI	77.531	77.459	77.454	77.426	77.427	76.301

[a]CV, coefficient of variation; *D*, fractal dimension (see footnote *b*); IJI, interspersion–juxtaposition index.
[b]$D = \{2 \times \text{logarithm(perimeter)}\}/\text{logarithm(area)}$.

(Table 3), ranging from 259 (3.05- and 6.10-m coverages) to 276 (20- and 30-m coverages). However, the rate of increase was much slower than without the dissolving process. For example, without the dissolving process the 30-m cell size had 238 more polygons than the original coverage, but when the dissolving process was used it had only 19 more polygons. Mean polygon size still decreased when grid cell sizes larger than 10 m were used, although the decrease was smaller than without dissolving. CV still increased with grid cell sizes over 10 m, and *D* decreased when the 30-m grid cell size was used. Edge density for the total landscape now decreased with larger cell sizes; IJI showed no definite trend.

In the youngest age-class, the number of polygons and edge density increased as grid cell size increased from 3.05 to 20 m and then decreased at 30 m. CV showed no definite trend, but it was highest at the 20-m grid cell size. *D* was lowest when the 30-m grid cell size was used, and IJI declined steadily across all cell sizes (Table 3). Similar trends were evident for the oldest age-class, with the exception of edge density, which decreased at grid cell sizes of 20 m and larger. IJI for both the youngest and oldest age-classes was greater than that for the total landscape.

Predicting changes in polygons as a result of data conversion

Individual polygons may be split into several smaller polygons, and the edges and shape may become more simplified because of data conversion. Thus while the polygon-level ANOVA presented earlier indicated significant changes (in terms of average polygon area, perimeter, and *D*) caused by data conversion, the ANOVA did not utilize the sum of the metrics for original polygons that were split during conversion. Regression equations (Table 4) developed to estimate the changes in individual polygons whose size was originally <10, 10–20 ha, and >20 ha show that only a small proportion of variation in the changes due to data conversion can be explained by either the original polygon metric or an indicator variable representing a grid cell size used in data conversion. Interestingly, we could not develop equations to predict the change in polygon area. Several of the equations, in fact, utilize only indicator variables to predict changes in polygon metrics and thus are of limited use for predicting changes due to data conversion from grid cell sizes other than those the indicator variables represent. Developing equations to predict changes in the original polygon metrics proved difficult for two reasons: (1) the size (and shape) of the original polygon may have been an influencing factor and (2) significant changes occurred primarily when larger grid cell sizes were used during data conversion.

Discussion

Our results indicate that scaling issues are important when data conversion processes are used, because most polygon-and

Table 3. Landscape-level indices based on GIS vegetation coverages with a dissolving process applied.

Index[a]	Original coverage	Grid cell size used in data conversion				
		3.05 m	6.10 m	10 m	20 m	30 m
Total landscape						
Number of polygons	257	259	259	260	276	276
Mean polygon size (ha)	12.301	12.208	12.214	12.167	11.450	11.454
Polygon CV (%)	105.629	106.377	106.341	106.832	111.512	110.957
D^b	1.315	1.314	1.314	1.314	1.314	1.308
Edge density (m/ha)	74.545	74.223	74.206	74.327	73.772	71.692
IJI	69.034	69.127	69.014	68.996	69.079	69.072
Youngest age-class (≤10 years old)						
Number of polygons	42	42	42	43	47	45
Mean polygon size (ha)	10.181	10.182	10.188	9.946	9.095	9.462
Polygon CV (%)	83.289	82.980	83.087	85.319	88.748	84.960
D^b	1.306	1.306	1.306	1.307	1.308	1.302
Edge density (m/ha)	17.745	17.810	17.798	17.813	17.813	17.318
IJI	78.752	79.019	78.828	78.757	78.626	78.085
Oldest age-class (≥140 years old)						
Number of polygons	90	90	90	90	94	91
Mean polygon size (ha)	11.711	11.719	11.735	11.726	11.185	11.599
Polygon CV (%)	81.614	81.832	81.897	81.846	85.977	82.901
D^b	1.319	1.316	1.316	1.316	1.314	1.307
Edge density (m/ha)	39.020	38.842	38.844	38.911	38.374	37.547
IJI	77.531	78.118	77.504	77.481	77.636	77.273

[a]CV, coefficient of variation; *D*, fractal dimension (see footnote *b*); IJI, interspersion–juxtaposition index.
[b]$D = \{2 \times \text{logarithm(perimeter)}\}/\text{logarithm(area)}$.

landscape-level indices are sensitive to changes in the spatial resolution used in the conversion process. Because of the dependence of polygon and landscape parameters on spatial resolution, the use of GIS data developed at a fine scale but generalized at large grid cell sizes may not be a solution to addressing questions regarding broad-scale issues. In this study, for example, if we assume that each polygon is an analysis unit, converting the data using the 30-m grid cell size and then using it for resource scheduling without an alleviation technique would cause the original vector polygons to be represented by many new (often small) vector polygons. Thus the original polygon could be assigned several different land management activities, even though the goal may have been to apply a single set of management activities to the area defined by the original polygon. Further, changes in the locations of polygon boundaries due to conversion may affect the representation of habitat corridors or other resources, such as the location of riparian areas. In this case study, some of the original riparian areas were thin and sinuous; after GIS data were converted at the 20- and 30-m grid cell sizes, many riparian areas became highly fragmented.

The use of large grid cells leads to coarser grained observations, reducing the observed heterogeneity even though finer scale heterogeneity is already integrated into the model grain; thus differences between components are obscured by smoothing, averaging, integration, and aggregation (King 1991). From an ecological point of view, conversion with larger grid sizes may misrepresent many ecological processes. It is unclear at these large grid cell sizes whether the landscape should be viewed as a highly fragmented landscape and (or) whether critical habitat corridors or connections are now separated by some artificially imposed aggregation process. From the land manager's point of view, the increase in numbers of polygons not only represents ecological inaccuracies, but also a gross inefficiency in planning and record keeping.

Conversion may also misrepresent the ecological function of edges identified at fine scales. At both analysis levels, perimeter (polygon level) and edge (landscape level) were sensitive to changes in grid cell size. It is unclear, from an ecological point of view, what effect the dissolving process had on the representation of edge. For example, changes in perimeter were significant at the polygon level when the dissolving process was not used, but not significant when the process was applied. At the landscape level, edge increased at larger grid cell sizes without the dissolving process. However, this increase was artificial because many edges were artifacts of the data conversion process. Once most of these artifacts were eliminated through the dissolving process, edge density decreased with increasing grid cell size.

Because of the bridging that occurred at narrow parts of polygons, many small polygons were created in the data

Table 4. Regression equations to predict changes in polygon perimeter and *D*.

Polygon metric	Regression equation	R^2
Without dissolving		
Perimeter		
<10 ha	243.016 − 32.619(log perimeter) + 47.427(D5)	0.07
10–20 ha	−8.014 − 96.568(D5)	0.23
>20 ha	−23.810 + 168.910(D5)	0.11
D		
<10 ha	−0.580 − 0.767(1/*D*) + 0.054(D5)	0.06
10–20 ha	−0.0007 − 0.006(D5)	0.26
>20 ha	−0.024 − 0.033(1/*D*) − 0.0024(*D*4)	0.31
With dissolving		
Perimeter		
<10 ha	25.062 + 102.841(D5)	0.19
>20 ha	−2536.170 + 313.611(log perimeter) + 267.080(D5)	0.31
D		
<10 ha	0.662 − 0.874(1/*D*) + 0.061(D5)	0.08
>20 ha	0.020 − 0.026(1/*D*) + 0.004(D5)	0.83

Note: D5 is a 0–1 variable indicating the use of 30 m grid cell size during data conversion. D4 is a 0–1 variable indicating the use of 20 m grid cell size during data conversion.

conversion process. In some cases small polygons became completely fragmented, but for larger polygons only the edges were broken off. Our expectation was that as grid cell size increased, fragmentation or bridging would increase. We found dramatic increases in CV, which measures the variation in polygon size relative to mean polygon size, when using either 20- or 30-m grid cell sizes. However, the effects of fragmentation and (or) bridging seem minor for grid cell sizes ranging from 3.05 to 10 m.

Our expectation was that the parameter describing texture (*D*) would be sensitive to scale (e.g., Benson and Mackenzie 1995). What we found, in fact, was that without the dissolving process, *D* seemed sensitive only to one grid cell size (20 m) at both the polygon and landscape levels. Thus for at least part of a range of scales (3.05 to 10 m), *D* stayed constant. The fact that *D* did not change shows its apparent self-similarity at these scales, with no dominance of different processes. The range of self-similarity may be only 3.05 to 10 m in this landscape; after 10 m, *D* behaved unpredictably when an alleviation process (e.g., dissolving) was not used. After use of the dissolving process, *D* seemed sensitive only to the 30-m grid cell size at the landscape level.

Prior to use of the dissolving process, IJI decreased as grid cell size increased. The significance of this change depends on the current state of the landscape and the goals we are hoping to fulfill. For instance, if our objective were to manage for large, contiguous blocks of old forest stands, we would want IJI to be low. If, on the other hand, we were to manage for a mixture of older and younger stands (which would provide better habitat for big game), we would want IJI to be relatively high. Therefore the IJI results seem to contradict the results obtained by measures of number of polygons, mean polygon size, and CV. The decreasing IJI implies that the landscape is becoming not more fragmented but more aggregated at larger grid cell sizes. However, it is the increase in total edge relative to the distribution of edge between any two age-classes that produces this anomalous result at the landscape level. Therefore, it is the artificial edge created within fragmented polygons that contributes to the decrease in IJI. After the dissolving process was used, IJI did not decrease with increasing grid cell size.

In our analysis, the youngest and oldest age-classes had higher IJI than the total landscape. These age-classes were apparently more interspersed than the age-classes between them. Forest cutting has been concentrated within the oldest age-classes, with harvests planned to be interspersed across the landscape. The intermediate age-classes may therefore have been left unmanaged (and clumped together) for many decades.

Our results indicate that if a vector–raster–vector data conversion process is used, grid cell sizes up to 10 m may give adequate data sets from GIS without significant changes in area, perimeter, or *D* at the polygon level. Changes in polygon boundary location were not examined; such changes may affect other goals, such as maintaining habitat corridors. When the dissolving process was not used in this case study, data sets converted at grid cell sizes of 20 m or larger differed significantly at the polygon level from the data set that best represented the original data, and they also seemed to differ at the landscape level in terms of number of polygons, mean polygon size, CV, edge density, and IJI. When a dissolving process was used, grid cell sizes up to 30 m seemed to provide adequate GIS data sets without significant changes in mean polygon area, polygon perimeter, and *D*. Further, at the landscape level, number of polygons, mean polygon size, CV, and edge density seemed to differ from the original data when 20- or 30-m grid cell sizes were used in the generalization process.

The results of this research may be valid only for our case study area. We selected the study area because it represents a "typical" forest and range landscape in eastern Oregon. Different results may occur if the GIS polygons are larger or smaller than the ones delineated in this study area. In modeling systems in a spatially explicit framework, it is essential to use spatial data that meet quality requirements set by the host organization. We have assumed in this study that the quality requirements were built into the data collection and encoding phases of building the database. Changes to the database through such operations as data extraction and conversion of format may move the data away from the initial quality standards. In an ideal situation, the data for a vector coverage would be exported for use in spatial analysis without alteration of the coordinates defining the geographic features. In reality, however, organizations are limited by the computer systems available to them.

References

Benson, B.J., and Mackenzie, M.D. 1995. Effects of sensor spatial resolution on landscape structure parameters. Landscape Ecol. **10**(2): 113–120.

Berenson, M.L., and Levine, D.M. 1983. Basic business statistics; concepts and applications. 2nd ed. Prentice-Hall, Englewood Cliffs, N.J.

Bradshaw, F.J. 1992. Quantifying edge effect and patch size for multiple-use silviculture—a discussion paper. For. Ecol. Manage. **48**: 249–264.

Bregt, A.K., Denneboom, J., Gesink, H.J., and Van Randen, Y. 1991. Determination of rasterizing error: a case study with the soil map of the Netherlands. Int. J. Geogr. Inf. Syst. **5**: 361–367.

Burrough, P.A. 1989. Matching spatial databases and quantitative models in land resource assessment. Soil Use Manage. **5**: 3–8.

Cartensen, L.W., Jr. 1989. A fractal analysis of cartographic generalization. Am. Cartogr. **16**: 181–189.

Carver, S.J., and Brunsden, C.F. 1994. Vector to raster conversion error and feature complexity: an empirical study using simulated data. Int. J. Geogr. Inf. Syst. **8**: 261–270.

Clarke, K.C. 1985. A comparative analysis of polygon to raster interpolation methods. Photogramm. Eng. Remote Sens. **51**: 575–582.

Clarke, K.C. 1990. Analytical and computer cartography. Prentice-Hall, Englewood Cliffs, N.J. pp. 186–197.

Dunn, C.P., and Loehle, C. 1988. Species–area parameter estimation: testing the null model of lack of relationship. J. Biogeogr. **15**: 721–728.

Dunn, C.P., Sharpe, D.M., Guntenspergen, G.R., Stearns, F., and Yang, Z. 1991. Methods for analyzing temporal changes in landscape pattern. *In* Quantitative methods in landscape ecology: an introduction. *Edited by* M.G. Turner and R.H. Gardner. Springer-Verlag, New York. pp. 173–198.

Environmental Systems Research Institute. 1994. ArcForest version 1.3. ESRI Canada Limited, Victoria, B.C.

Forest Ecosystem Management Assessment Team. 1993. Forest ecosystem management: an ecological, economic, and social assessment. U.S. Department of Agriculture, U.S. Department of the Interior, et al., Portland, Oreg.

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

Franklin, J.F., and Forman, R.T.T. 1987. Creating landscape patterns by forest cutting: ecological consequences and principles. Landscape Ecol. **1**: 5–18.

Gates, J.E., and Gysel, L.W. 1978. Avian nest dispersion and fledgling success in field–forest ecotones. Ecology, **59**: 871–883.

Goodchild, M.F. 1980. Fractals and the accuracy of geographical measures. Math. Geol. **12**: 85–98.

Haining, R., and Arbia, G. 1993. Error propagation through map operations. Technometrics, **35**: 293–305.

Hansen, A.J., Garman, S.L., Lee, P., and Horvath, E. 1993. Do edge effects influence tree growth rates in Douglas-fir plantations? Northwest Sci. **67**: 112–116.

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

Hunter, M.L., Jr. 1990. Wildlife, forests, and forestry; principles of managing forests for biological diversity. Prentice-Hall, Englewood Cliffs, N.J.

King, A.W. 1991. Translating models across scales in the landscape. *In* Quantitative methods in landscape ecology: an introduction. *Edited by* M.G. Turner and R.H. Gardner. Springer-Verlag, New York. pp. 479–518.

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.

Li, H., and Reynolds, J.F. 1994. A new contagion index to quantify spatial patterns of landscapes. Landscape Ecol. **8**(3): 155–162.

Loh, D.K., and Rykiel, E.J., Jr. 1992. Integrating resource management systems: coupling expert systems with data-base management and geographic information systems. Environon. Manage. **16**: 167–177.

McGarigal, K., and Marks, B.J. 1993. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure, version 1.0. Department of Forest Science, Oregon State University, Corvallis.

McIntyre, N.E. 1995. Effects of patch size on avian diversity. Landscape Ecol. **10**(2): 85–99.

Milne, B.T. 1991. Lessons from applying fractal models to landscape patterns. *In* Quantitative methods in landscape ecology: an introduction. *Edited by* M.G. Turner and R.H. Gardner. Springer-Verlag, New York. pp. 199–238.

Ripple, W.J., Bradshaw, G.A., and Spies, T.A. 1991. Measuring forest landscape patterns in the Cascade Range of Oregon, USA. Biol. Conserv. **57**: 73–88.

Schuster, E.G., Leefers, L.A., and Thompson, J.E. 1993. A guide to computer-based analytical tools for implementing forest plans. USDA For. Serv. Gen. Tech. Rep. INT-296.

Sessions, J., and Sessions, J.B. 1993. SNAP II+ user's guide. Department of Forest Engineering, Oregon State University, Corvallis.

Tomlin, C.D. 1990. Geographic information systems and cartographic modeling. Prentice-Hall, Englewood Cliffs, N.J.

USDA Forest Service. 1992. LTPlus (Line Trace Plus). Version 2.36X. USDA Forest Service, Geometronics Service Center, Salt Lake City, Utah.

Valenzuela, C.R., and Baumgardner, M.F. 1990. Selection of appropriate cell sizes for thematic maps. ITC J. **4**: 219–224.

Veregin, H. 1989. A review of error models for vector to raster conversion. Oper. Geogr. **4**: 11–15.

Walsh, S.J., Lightfoot, D.R., and Butler, D.R. 1987. Recognition and assessment of error in geographic information systems. Photogramm. Eng. Remote Sens. **53**: 1423–1430.

Wehde, M. 1982. Grid cell size in relation to errors in maps and inventories produced by computerized map processing. Photogramm. Eng. Remote Sens. **48**: 1289–1298.

Wiens, J.A. 1989. The ecology of bird communities. Vol. 1. Foundations and patterns. Cambridge Studies in Ecology, Cambridge University Press, Cambridge, Great Britain.