A Comparison of Vector and Raster GIS Methods for Calculating Landscape Metrics Used in Environmental Assessments

Timothy G. Wade, James D. Wickham, Maliha S. Nash, Anne C. Neale, Kurt H. Riitters, and K. Bruce Jones

Abstract

GIS-based measurements that combine native raster and native vector data are commonly used in environmental assessments. Most of these measurements can be calculated using either raster or vector data formats and processing methods. Raster processes are more commonly used because they can be significantly faster computationally than vector, but error is introduced in converting vector data to raster. This conversion error has been widely studied and quantified, but the impact on environmental assessment results has not been investigated. We examined four GIS-based measurements commonly used in environmental assessments for approximately 1000 watersheds in the state of Maryland and Washington, D.C. Each metric was calculated using vector and raster methods, and estimated values were compared using a paired t-test, Spearman rank correlation, and cluster analyses. Paired t-tests were used to determine the statistical significance of quantitative differences between methods, and Spearman rank correlation and cluster analyses were used to evaluate the impact of the differences on environmental assessments. Paired t-test results indicated significant quantitative differences between methods for three of the four metrics. However, Spearman ranks and cluster analyses indicated that the quantitative differences would not affect environmental assessment results. Spearman rank correlations between vector and raster values were greater than 0.98 for all comparisons. Cluster analyses resulted in identical assignment for 88 percent to over 98 percent of watersheds analyzed among vector and various raster methods.

Introduction

Over the last decade there has been a noticeable expansion in the development of raster-based land-cover data sets (e.g., Loveland *et al.*, 1991; Scott *et al.*, 1993; Dobson *et al.*, 1995;

T.G. Wade and J.D. Wickham are with the U.S. Environmental Protection Agency, National Exposure Research Laboratory, E243-05, Research Triangle Park, NC 27711 (wade.timothy@epa.gov; wickham.james@epa.gov).

M.S. Nash, A.C. Neale, and K.B. Jones are with the U.S. Environmental Protection Agency, National Exposure Research Laboratory, Box 93478, Las Vegas, NV 89193-3478 (nash.maliha@epa.gov; neale.anne@epa.gov; jones. bruce@epa.gov).

K.H. Riitters is with the U.S. Department of Agriculture Forest Service, Southern Research Station, Box 12254, Research Triangle Park, NC 27709 (kriitters@fs.fed.us).

Loveland *et al.*, 1996; Scott and Jennings, 1998; Vogelmann *et al.*, 2001). When combined with other geographic data in a geographic information system (GIS), it is possible to generate suites of measurements for environmental applications (e.g., Jones *et al.*, 1997; Wickham *et al.*, 1999; Mehaffey *et al.*, 2001).

Use of GIS is now commonplace for generation of metrics to assess the environmental quality of an area. Land cover, roads, streams, and other spatial data can be combined to generate many measurements of environmental quality (Forman and Alexander, 1998). Simple metrics such as road density are useful for habitat assessments (Lyon, 1979; Lyon, 1983; Thiel, 1985; Cox *et al.*, 1994; Forman and Alexander, 1998), and are also important for estimates of impervious surface (Arnold and Gibbons, 1996). Overlays of land cover or roads with streams can provide useful measurements of riparian habitat and water quality (Peterjohn and Correll, 1984; Osborne and Kovacic, 1993; Jones *et al.*, 1997; Wickham *et al.*, 2002).

The native format of several data layers used in metric generation, such as roads, streams, and soils, is typically vector (USGS, 1989; Bliss and Reybold, 1989), while land cover and slope are more typically raster (Loveland *et al.*, 1991; Scott *et al.*, 1993; Dobson *et al.*, 1995; Scott and Jennings, 1998; Loveland *et al.*, 1999; Vogelmann *et al.*, 2001). Derivative measurements that combine raster and vector data require conversion of all data to either the vector or raster domain for processing.

Historically, it has generally been accepted that raster calculations were faster, while vector methods gave higher accuracy (Kennedy and Meyers, 1977). In the past, landscape-level and larger studies used raster methods because computer hardware and software limitations prohibited use of vector methods. The rule of "raster is faster, vector is better" still holds true with the possible exception of very fine-scale (small cell size) raster processing.

Errors related to integration of remotely sensed data into GIS were reviewed by Lunetta *et al.* (1991), while Walsh *et al.* (1987) quantified operational error resulting from overlaying remotely sensed data in a GIS. Several studies have documented the errors associated with the vector to raster or raster to vector conversion process (Clarke, 1985; Veregin, 1989). Much of this work focused on area (polygon) conversions and

the affects of polygon size and shape and raster cell size (Wehde, 1982; Carver and Brunsdon, 1994; Congalton, 1997). Linear conversion errors are less well studied, but Theobald (2000) found length measurement errors increased with increasing raster resolution. However, little work has been done on how raster and vector methods, and their associated conversion errors, affect landscape metrics or whether those errors are large enough to impact environmental assessments. Bettinger *et al.* (1996) showed how polygon conversion affected some forest patch metrics, and Wickham *et al.* (1996) showed that some measurements can be dependent on data format.

Many landscape metrics are used in environmental assessments (Jones *et al.*, 1996; Jones *et al.*, 1997; O'Neill *et al.*, 1997; Wickham *et al.*, 1999; Jones *et al.*, 2001; Mehaffey *et al.*, 2001; Wickham *et al.*, 2002). Assessments often cover large areas, and metrics are usually calculated using raster methods. The objective of this paper is to determine whether the differences in metric values generated from vector and raster GIS methods might affect assessments based on the metric values. We calculated four commonly used landscape measurements by watershed across the state of Maryland and compared results by method. Comparisons were made with paired t-tests, which determined the statistical significance of the differences between methods, and Spearman rank correlations, absolute differences, and cluster analyses, all of which assessed the importance of the differences to environmental assessments.

Methods

Four GIS-based measurements of environmental quality were chosen based on their frequent use and mix of vector and raster data inputs. They were (1) land-cover proportions by watershed (LCP), (2) land-cover proportion adjacent to streams (LCS), (3) land-cover proportion in riparian zones (LCRZ), and (4) road density in riparian zones (RDST).

We reported metric values by 12-digit watersheds for the state of Maryland (Figure 1), which supplied 1041 assessment units (watersheds) at a scale that would be used by environmental managers. Maryland was chosen for its high degree of diversity. The state covers four physiographic provinces (Coastal Plain, Piedmont, Valley-and-Ridge, Appalachian Plateau), and the watersheds provided a full gradient of homogeneous to heterogeneous land cover, so our results should be applicable in other areas.

Native vector data included watershed boundaries, roads, and streams. Watersheds were obtained from the Maryland Department of Natural Resources (MD DNR). For roads, we used the Wessex Streets (version 7.0), and we used the USGS Digital Line Graph (DLG) hydrography data for streams. Hydrography date were not available for one quadrangle in western Maryland (McHenry) or any of the Washington, D.C. area. As a result, only 1015 watersheds were available for metrics that required stream data.

The National Land Cover Dataset (NLCD; Vogelmann *et al.*, 2001) was used to calculate all land-cover proportions. These data were derived from Landsat TM and other ancillary data (Vogelmann *et al.*, 2001) and closely follow an Anderson Level II classification (Anderson *et al.*, 1976). We reclassified the 21 class legend into six classes: forest, developed (urban), agriculture, water, wetland, and barren. Land-cover proportions were based on terrestrial area only (i.e., the water class was ignored). These data were obtained in raster format, with a 30-meter spatial resolution.

Metrics were generated using Arc/Info workstation version 8.1 with the Grid extension (ESRI, 2002). The software used the dominant method in converting vector to raster, so each cell was assigned the value of the polygon that occupied the largest amount of cell area. All raster datasets were created such that cell boundaries were coincident with land-cover cells. To convert raster to vector, polygon boundaries followed cell edges. The minimum polygon size was a single pixel, or 900 square meters. No polygons were eliminated or aggregated.

Proportion of Land Cover (LCP)

The proportion of land-cover types is a commonly used measurement of overall environmental quality at regional scales (Jones *et al.*, 1997; Paul *et al.*, 2002). It is routinely included in software packages that provide an array of landscape measurements (Baker and Cai, 1992; McGarigal *et al.*, 1995). In vector processing, the land cover was converted to polygons, and intersected with the vector watershed boundaries. Proportions of each land-cover type were then calculated for the watersheds. In raster processing, the land cover was left in its native raster format and overlaid with a rasterized watershed boundaries, then proportions were calculated.

Land Cover Adjacent to Streams (LCS)

The extent to which other land-cover types replace forest in the riparian zone can be used as a local (Peterjohn and Correll, 1984), regional (Lowrance *et al.*, 1997), and national
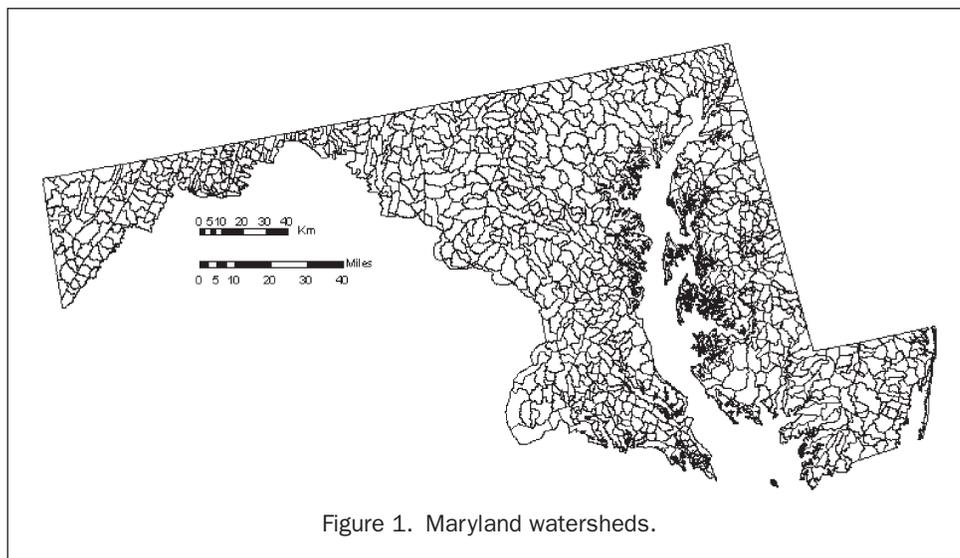

Figure 1. Maryland watersheds.

(Swift, 1984) measurement of water quality and habitat impacts. Land cover adjacent to streams utilized both vector (streams) and raster (land-cover) data. In the vector processing, land cover was again converted to polygons and overlaid with streams. Land-cover proportions were calculated as the ratio of stream length passing through each cover type to total stream length in the watershed. For raster processing, streams were rasterized using a 30-meter cell size, and overlaid with the land-cover grid. Proportions were calculated as the ratio of the number of cells of each cover type to total number of stream cells in the watershed.

Riparian Zone Land Cover (LCRZ)

This metric is a combination of LCP and LCS, because it computed the proportion of land cover within a buffer zone around streams. In the vector processing, streams were buffered by 45, 75, 105, 135, and 165 meters on each side, resulting in riparian corridors of 90, 150, 210, 270, and 330 meters, respectively. These buffers were then intersected with polygon land cover. The land-cover proportions were calculated as the ratio of area of each cover type within the buffer zone to total area of the buffer zone in the watershed.

Two raster methods, both based on the rasterized streams used in the LCS metric, were used to create buffers for the LCRZ measurements. In addition, a hybrid vector/raster method was used to generate buffers. The first raster method (hereafter EXP) "grew" the rasterized streams by a specified number of cells by adding cells that were adjacent to a rasterized stream cell in either a cardinal or diagonal direction. This process was performed with one, two, three, four, and five cell expansions, which resulted in total swath widths of 90, 150, 210, 270, and 330 meters, the same as for the vector buffers. This is the easiest and most commonly used method to generate raster buffers.

The second raster method (hereafter EUC) used Euclidean distance to locate cells with centers within 30, 60, 90, 120, and 150 meters of a rasterized stream cell center. Because the streams were represented by a 30-meter cell, this also resulted in total swath widths of 90, 150, 210, 270, and 330 meters. The buffers were then overlaid with the raster land cover to determine proportions.

The hybrid method used buffers created as a vector process from vector streams. The buffers were then rasterized and overlaid with raster land cover to compute proportions.

Roads Density in Riparian Zones (RDST)

RDST measured length of roads within 45, 75, 105, 135, and 165 meters of a stream. In vector processing, streams were buffered as in the riparian zone metric and overlayed with roads. The RDST value was the total road length within the buffer divided by total stream length in the watershed, multiplied by 1000 (meters of road per kilometer of stream). Raster processing required conversion of both roads and streams from vector to raster format with a 30-meter cell size. Streams were then buffered using the EXP and EUC raster methods and overlaid with roads. The number of road cells within the stream buffer was divided by total number of stream cells in the watershed and multiplied by 1000 to calculate RDST. The hybrid process used in the LCRZ metric was also used here. Vector streams were buffered, and the buffers were rasterized and overlaid with rasterized roads to calculate road density near streams.

Statistical Analyses

Comparison of metric results between vector, raster, and hybrid methods included paired t-tests, average absolute difference (AAD), Spearman rank correlation, and cluster analyses. The objective of the t-tests was to determine whether differences in metric values between methods were statistically significant. The objectives of the Spearman r, AAD, and cluster analyses were to assess the importance of any statistically significant differences to an assessment.

A two-tailed, paired t-test ($\alpha = 0.05$) was performed separately on each of the five land-cover classes for each metric except RDST, which did not use land cover in its calculation. The paired t-test null hypothesis was that the mean difference between the two methods was 0. The paired t-test uses differences between cases in its calculation, so even a small bias in either the positive or negative in the case differences may result in the rejection of the null hypothesis. Because of this, we also report the AAD, or |vector − raster|, averaged across all observations.

Many environmental assessments assign units to groups based on metric scores (Jones *et al.*, 1997; O'Neill *et al.*, 1997; Wickham *et al.*, 1999). T-test results may identify significantly different estimates, but cannot indicate whether such differences would result in assignment to a different group. Spearman r and cluster analyses were used to determine whether differences in metric values between raster and vector formats changed group assignment. High values of Spearman r between methods would indicate that the rank ordering of watersheds did not change despite differences discerned by the paired t-tests. Consequently, grouping (e.g., quantiles) watersheds would not be impacted by processing method.

Cluster analyses only included LCP, LCS, and the 165-meter buffer for LCRZ and RDST. The results from all buffer sizes for LCRZ and RDST were highly correlated, so we chose to use the 165-meter buffer as an intermediate scale between stream bank (LCS) and entire watershed (LCP). Metrics for each method were converted to principal components (PC) so that uncorrelated data were used in the cluster analyses. Each method produced four PCs with eigenvalues greater than one that explained at least 92 percent of the variance (cumulative) in the original data, and eigenvector loadings that were consistent across all methods. The PCs were output standardized to a mean of zero and unit variance, because cluster analysis is sensitive to differences in magnitude in the input data. The SAS procedure FASTCLUS was used with the maximum number of clusters set to ten. The first iteration of the clustering was taken as the final output, so that seed selection did not influence final cluster assignment. Ten was chosen for the maximum number of clusters because it was assumed that inspection of more groups would make it difficult to incorporate the information into management-related assessments.

There is negligible thematic error in converting land cover from raster to vector in Grid because polygon boundaries coincide with cell boundaries, but this is not true when converting native vector data to a raster format (Veregin, 1989; Congalton, 1997). Therefore, vector results were treated as "truth," or the reference to which raster and hybrid results were compared in our analyses.

Results

Both methods gave very similar results for the LCP metric. In the paired t-test, no significant difference was found between methods for any land-cover type except wetland (Table 1, Figure 2). AADs were less than 0.02 percentage points for all land-cover classes. Spearman r scores were 0.999 or higher for all land-cover types. The LCS paired t-test found significant differences between methods for all land-cover classes, but the AADs were less than 0.9 for all classes. Spearman correlations were 0.997 or higher (Table 1, Figure 2).

The LCRZ paired t-test showed significant differences between vector and both raster methods for nearly all land-cover classes and buffer sizes. Exceptions were the forest class 45-meter buffer (EXP method only) and the 45- and 75-meter buffers (EUC method only). Results were also significantly

TABLE 1. STATISTICAL RESULTS FOR RASTER VS. VECTOR LCP
AND LCS (AAD = AVERAGE ABSOLUTE DIFFERENCE)

Metric	Land Cover	t Value	AAD	Spearman r
LCP	Developed	**−0.39**	0.007	1.000
	Forest	**0.68**	0.015	1.000
	Agriculture	**0.74**	0.015	1.000
	Barren	**0.42**	0.002	0.999
	Wetland	−2.26	0.006	1.000
LCS	Developed	2.05	0.243	0.997
	Forest	−5.42	0.774	0.999
	Agriculture	−21.34	0.891	0.999
	Barren	−3.48	0.043	0.998
	Wetland	21.53	0.089	0.999

Bold indicates no significant difference, $\alpha = 0.05$.

different between vector and hybrid results for smaller buffers, but no significant differences were found for developed or barren at 105 meters and larger or wetland at 135 and 165 meters. Furthermore, there was no significant difference between vector and hybrid methods for any land-cover type in the 165-meter buffer. There was no discernible trend between vector and EXP results across buffer widths, but vector/hybrid comparisons became increasingly more similar as buffer size increased. Vector versus EUC results also became more similar except for the 135-meter buffer. AADs were lowest for the hybrid method for all buffers and land cover. EUC AADs were consistently next lowest, and EXP AADs were always highest. Spearman correlations were lowest for barren for all methods, ranging from 0.980 to 0.997. For all other land-cover types, the correlations were 0.998 or higher (Table 2, Figure 2) .

The paired t-test reported significant differences between vector and all three methods for the RDST metric for all buffer sizes. AADs were between 7 and 121 (meters of road/km of stream), with the hybrid method having the lowest values. Spearman r results were between 0.986 and 0.996 for all methods and buffer sizes (Table 2).

Cluster analyses produced nearly identical groupings for vector, EUC, and hybrid methods. Pairwise comparison of vector versus EUC and vector versus hybrid show agreement for 1000 of 1015 watersheds (98.5 percent), and comparison of the vector and EXP method agreed on 88.2 percent of watersheds (Table 3).

Raster (EXP) processing was significantly faster than vector processing. To generate all four metrics (including all necessary format conversions) took slightly over 10 minutes using EXP methods to almost 15 hours for vector methods, or over 80 times faster for the EXP method. EUC processing time for LCRZ and RDST was similar to EXP. The processing time for LCRZ and RDST with hybrid methods was five hours. In addition, vector files were significantly larger than raster files. For example, the reclassified land cover for the state of Maryland was 430 MB as a vector Arc/Info coverage, and contained over 900,000 polygons with an average size of almost 35,000 square meters. The land cover in grid format was only 11.4 MB. In general, vector files were 35 to 55 times larger than their raster counterparts.

Discussion

ArcInfo uses the dominant-area method to convert polygons to raster. Along watershed boundaries, land-cover cells are assigned entirely to the watershed that contains the most area in the cell. Vector processing allows the cell to be split and parts assigned to different watersheds. This edge effect has little impact when the reporting unit is significantly larger than the
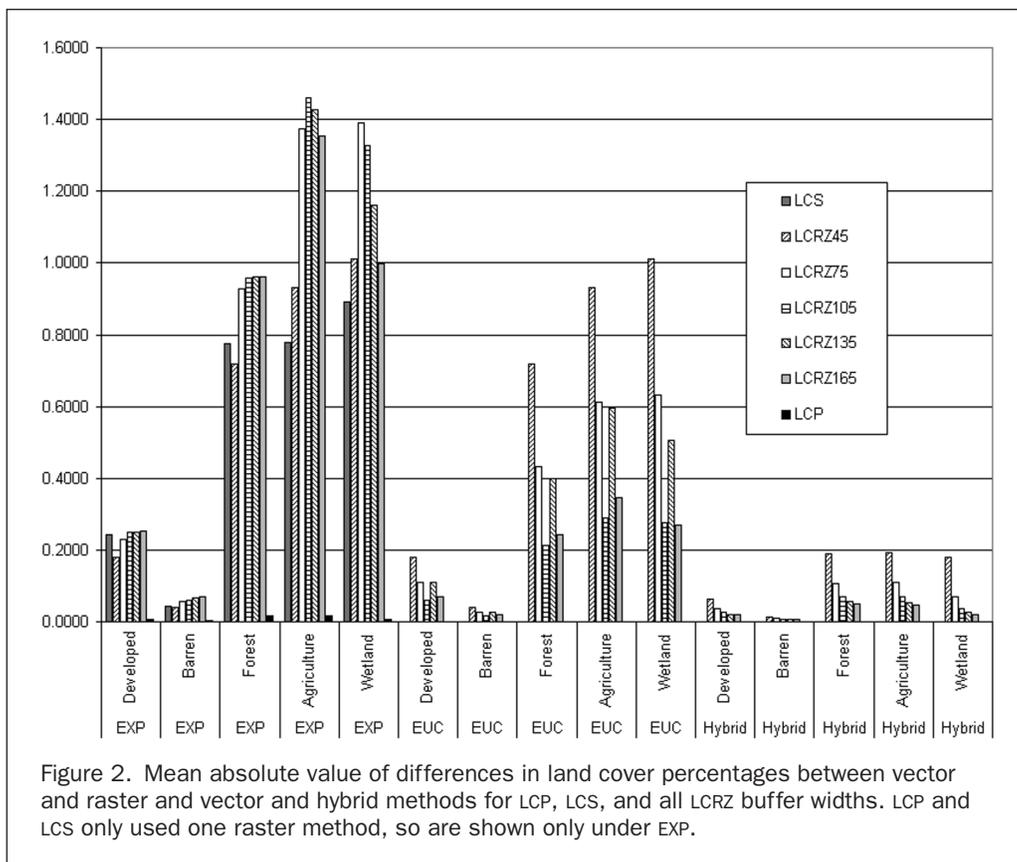

Figure 2. Mean absolute value of differences in land cover percentages between vector and raster and vector and hybrid methods for LCP, LCS, and all LCRZ buffer widths. LCP and LCS only used one raster method, so are shown only under EXP.

TABLE 2. STATISTICAL RESULTS FOR LCRZ AND RDST (AAD = AVERAGE ABSOLUTE DIFFERENCE, EXP = VECTOR/RASTER EXPAND COMPARISON, EUC = VECTOR/RASTER EUCLIDEAN COMPARISON, AND H = VECTOR/HYBRID COMPARISON)

		t Value			AAD			Spearman r		
Metric	Land Cover	EXP	EUC	H	EXP	EUC	H	EXP	EUC	H
LCRZ-45m	Developed	−4.10	−4.10	3.35	0.18	0.18	0.06	0.998	0.998	0.999
LCRZ-75m		−8.54	−6.72	4.62	0.23	0.11	0.04	0.998	0.999	1.000
LCRZ-105m		−10.48	−7.00	**1.53**	0.25	0.06	0.03	0.998	1.000	1.000
LCRZ-135m		−10.70	−9.68	**−1.48**	0.25	0.11	0.02	0.998	1.000	1.000
LCRZ-165m		−10.71	−8.71	**−0.68**	0.25	0.07	0.02	0.998	1.000	1.000
LCRZ-45m	Forest	**−1.91**	**−1.91**	**−0.53**	0.72	0.72	0.19	0.999	0.999	1.000
LCRZ-75m		2.01	**0.41**	−3.45	0.93	0.43	0.11	0.999	1.000	1.000
LCRZ-105m		6.62	2.40	−4.05	0.96	0.21	0.07	0.999	1.000	1.000
LCRZ-135m		9.92	8.60	−1.97	0.96	0.40	0.06	0.999	1.000	1.000
LCRZ-165m		11.13	9.28	**−1.29**	0.96	0.24	0.05	0.999	1.000	1.000
LCRZ-45m	Agriculture	−32.63	−32.63	18.23	0.93	0.93	0.19	0.999	0.999	1.000
LCRZ-75m		−35.73	−34.06	11.90	1.37	0.61	0.11	0.999	1.000	1.000
LCRZ-105m		−34.75	−29.78	7.43	1.46	0.29	0.07	0.999	1.000	1.000
LCRZ-135m		−32.93	−33.10	3.84	1.43	0.60	0.05	0.999	1.000	1.000
LCRZ-165m		−29.56	−29.05	**1.58**	1.35	0.34	0.05	0.999	1.000	1.000
LCRZ-45m	Barren	−5.75	−5.75	3.43	0.04	0.04	0.01	0.993	0.993	0.980
LCRZ-75m		−7.12	−6.16	3.49	0.06	0.03	0.01	0.991	0.988	0.981
LCRZ-105m		−7.56	−5.60	**1.62**	0.06	0.02	0.01	0.993	0.993	0.991
LCRZ-135m		−8.16	−7.64	**1.33**	0.07	0.03	0.01	0.995	0.997	0.995
LCRZ-165m		−8.47	−7.42	**0.69**	0.07	0.02	0.01	0.992	0.993	0.992
LCRZ-45m	Wetland	30.33	30.33	−21.03	1.01	1.01	0.18	0.999	0.999	1.000
LCRZ-75m		33.00	32.63	−17.08	1.39	0.63	0.07	0.999	1.000	1.000
LCRZ-105m		32.43	31.42	−7.09	1.33	0.28	0.04	0.999	1.000	1.000
LCRZ-135m		31.02	31.57	**−1.75**	1.16	0.51	0.03	0.999	1.000	1.000
LCRZ-165m		29.03	29.47	**−0.24**	1.00	0.027	0.02	0.999	1.000	1.000
RDST-45m	NA	−45.06	−45.06	−5.07	39.92	39.92	7.33	0.986	0.986	0.992
RDST-75m		−46.82	−41.76	−4.96	63.45	30.29	10.78	0.989	0.994	0.995
RDST-105m		−42.02	−26.67	−5.10	85.03	20.99	14.25	0.989	0.996	0.996
RDST-135m		−47.90	−42.40	−3.56	103.96	44.43	18.33	0.990	0.996	0.995
RDST-165m		−47.78	−30.15	−2.97	120.97	34.88	22.23	0.988	0.995	0.995

Bold indicates no significant difference, $\alpha = 0.05$.

TABLE 3. CLUSTER ANALYSES WATERSHED COINCIDENCE

Cluster	Vector	Hybrid	EUC	EXP
1	196	192	191	125
2	2	2	2	2
3	1	1	1	1
4	306	304	304	290
5	316	312	313	310
6	67	63	63	62
7	27	27	27	24
8	22	22	22	22
9	35	34	34	35
10	43	43	43	24
Total	1015	1000 (98.5%)	1000 (98.5%)	895 (88.2%)

pixel size and perimeter to area is low, as in LCP, but has much more influence on metrics like LCRZ. The LCP vector and raster results were very similar with the exception of wetlands. Wetland is a relatively rare class in Maryland, mostly found near the coast, which is also the watershed boundary and has a complex shape. The high ratio of wetland on a boundary edge to total wetland magnified the edge effect error, and the paired t-test reported a significant difference between methods.

In the conversion of linear vector data like streams to a raster format, any cell that is intersected by any length of one or more streams is considered a stream cell in the output. There is no distinction between a cell that contains a very short length of stream and one that contains a long stream length (Figure 3). This phenomenon accounts for all of the
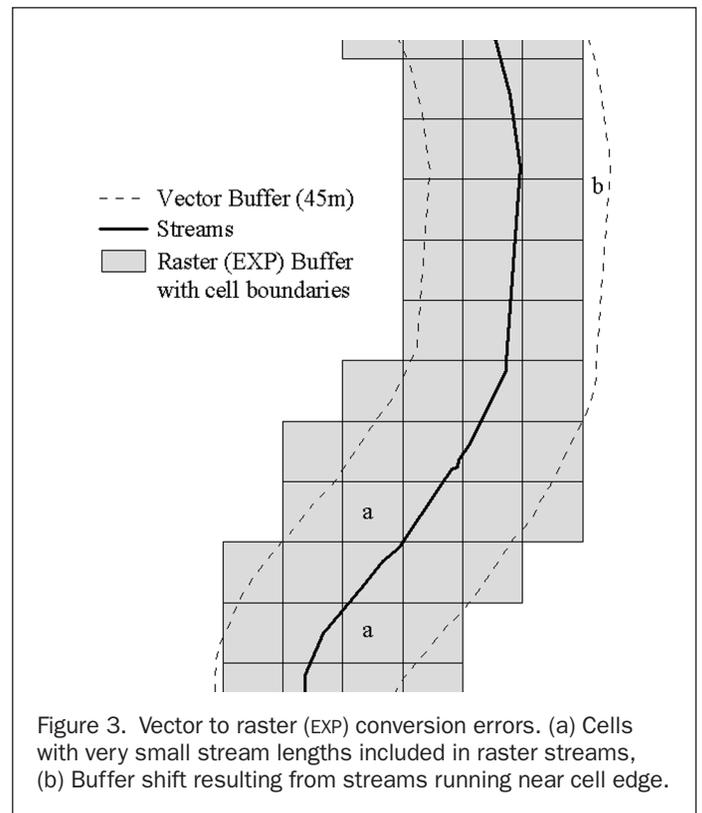

Figure 3. Vector to raster (EXP) conversion errors. (a) Cells with very small stream lengths included in raster streams, (b) Buffer shift resulting from streams running near cell edge.

Buffer Width	EXP	EUC	H
90 meters	23.78	23.78	0.34
150 meters	22.23	10.09	0.21
210 meters	20.81	4.01	0.15
270 meters	19.58	7.94	0.13
330 meters	18.23	4.47	0.11

differences in LCS between methods and can be mitigated by decreasing the cell size.

The LCRZ and RDST raster results were affected by the buffers used for these metrics. Despite the attempt to make buffers as similar as possible, total area contained within the vector buffer area was much smaller than either raster method for every watershed and all buffer widths. This difference was not related to data conversion but was simply a limitation of raster processing. Raster buffers became more similar in size to vector buffers as buffer width increased, but EUC converged much more quickly than EXP (Table 4). The average vector and hybrid areas per watershed were very similar for all buffers (Table 4), but edge effects were large enough to show significant differences in the paired t-test for small buffers. As buffer size increased, perimeter to area decreased, and hybrid values converged with vector values (Table 2, Figure 2).

The larger area in the raster buffer areas altered the proportions of land cover contained within the buffers. The 45-meter EXP and the 45- and 75-meter EUC paired t-test showed no significant difference from vector values for proportion of forest. This indicates that the forest proportions were similar in the smaller vector buffer and larger raster buffers, but other land-cover types were not. The differences between vector and EUC buffer sizes and associated AAD values were highly correlated for all land-cover types (r^2 between 0.92 and 0.98), indicating that most of the differences in metric values were due to buffer creation methodology.

The area contained in the buffers is particularly important when the metric has a known area of influence. As a hypothetical example, if it is known that habitat for a certain species must be within 60 meters of a stream, an EXP buffer will include more area and possibly overestimate potential habitat.

Stream location can also affect buffer creation. A stream positioned near a raster cell edge is treated as if it ran down the center when rasterizing, essentially shifting the buffer associated with the stream (Figure 3). This shift may alter the composition of the land cover within the buffer, with changes more likely in small buffers and heterogeneous landscapes (many small patches) than larger buffers and homogeneous areas.

While t-tests often indicated statistically different values, Spearman r values for all metrics indicated that all watersheds would show nearly identical rank orders regardless of the method. The lowest Spearman r values were 0.981. Thus, an assessment based on groupings of watersheds by rank orders would not be affected by the method used to calculate the metrics.

Cluster analyses were largely consistent with the Spearman r results. Nearly all (over 98 percent; Table 3) watersheds were assigned to the same group across vector, Euclidean (EUC), and hybrid (H) methods. Only the expand (EXP) method produced results that were somewhat inconsistent, assigning only 895 of the 1015 (88 percent; Table 3) watersheds to the same groups.

In the past, large-scale environmental assessments based on GIS-generated metrics nearly always used raster methods, usually EXP. Even with the improvement in computer hardware and software, raster methods are still predominantly used for these types of studies, because they are substantially faster. Software limitations (Arc/Info imposes a file size limit of 2.1 GB) are more likely to prohibit the use of vector methods for large-scale studies because vector files tend to be much larger than comparable raster files.

Conclusion

Paired t-tests showed statistically significant differences for most of the metrics presented here. However, the t-test is fairly sensitive, and the Spearman r, AAD, and cluster analyses all indicate that differences between methods for most metrics may be insignificant to an assessment.

Unfortunately, there is no definitive answer as to what method should be used. For assessments based on rankings or groups, our results indicate that any of the methods are sufficient, with the EXP option being the least desirable. If highly accurate individual observations are required, vector methods should be employed when possible. Assessment needs will determine which processing method is appropriate for a given metric. When the reporting unit is large relative to the pixel size (e.g., LCP), the method will have little or no impact on the assessment and the raster method is preferred for its greater efficiency. For metrics that use a mix of raster and linear vector data (LCS, LCRZ, RDST, etc.), Spearman ranks, AAD, and cluster analyses all indicate that there are very small differences between methods for most metrics. For many assessments, the faster raster or hybrid methods will provide adequate results, especially when buffer size is large.

Metrics that require a buffer generated from linear features should avoid the EXP raster method because the expansion distance is not indicative of true buffer width. The EUC method provides better results with no increase in processing time. When the metric has a known area of influence, a vector or hybrid method should be used so that the buffer is accurately defined. Whichever method is used to generate buffers should be documented (see Riitters *et al.*, 1996), because the area included can vary substantially. Metrics that are entirely based on native vector data, such as RDST, have the largest differences between methods, and may require vector or hybrid methods.

References

Anderson, J.F., E.E. Hardy, J.T. Roach, and R.E. Witmer, 1976. *A Land Use and Land Cover Classification for Use with Remote Sensor Data*, U.S. Geological Survey Professional Paper 964, U.S. Government Printing Office, Washington, D.C., 28 p.

Arnold, C.L., and C.J. Gibbons, 1996. Impervious surface: The emergence of a key environmental indicator, *Journal of the American Planning Association*, 62:244–252.

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

Bettinger, P., G.A. Bradshaw, and G.W. Weaver, 1996. Effects of geographic information system vector-raster-vector data conversion on landscape indices, *Canadian Journal of Forest Research*, 26:1416–1425.

Bliss, N.B., and W.U. Reybold, 1989. Small-scale digital soil maps for interpreting natural resources, *Journal of Soil and Water Conservation*, 44:30–34.

Carver, S.J., and C.F. Brunsdon, 1994. Vector to raster conversion error and feature complexity: An empirical study using simulated data, *International Journal of Geographic Information Systems*, 8(3):261–270.

Clarke, K.C., 1985. A comparative analysis of polygon to raster interpolation methods, *Photogrammetric Engineering & Remote Sensing*, 51(5):575–582.

Congalton, R.G., 1997. Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion, *Photogrammetric Engineering & Remote Sensing* 63(4):425–434.

Cox, J., R. Kautz, M. Maclaughlin, and T. Gilbert, 1994. *Closing the Gaps in Florida's Wildlife Habitat Conservation System*, Office of Environmental Services, Florida Game and Fresh Water Fish Commission, Tallahassee, Florida, 239 p.

Dobson, J.E., E.A. Bright, R.L. Ferguston, D.W. Field, L.L. Wood, K.D. Haddad, K. Iredale, J.R. Jensen, V.V. Klemas, R.J. Orth, and J.P. Thomas, 1995. NOAA *Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation*, NOAA Technical Report NMFS 123, National Oceanic and Atmospheric Agency, Seattle, Washington, 139 p.

ESRI, 2002. *ArcGIS B ArcInfo*, ESRI, Redlands, California, URL: http://www.esri.com/software/arcgis/arcinfo/index.html, last accessed 05 August 2003.

Forman, R.T.T., and L.E. Alexander, 1998. Roads and their major ecological effects, *Annual Review of Ecology and Systematics*, 29:207–231.

Jones, K.B., J. Walker, K.H. Riitters, J.D. Wickham, and C. Nicoll, 1996. Indicators of Landscape Integrity, *Indicators of Catchment Health: A Technical Perspective* (J. Walker and D.J. Reuter, editors), CSIRO, Melbourne, Australia, pp. 155–168.

Jones, K.B., K.H. Riitters, J.D. Wickham, R.D. Tankersley, Jr., R.V. O'Neill, D.J. Chaloud, E.R. Smith, and A.C. Neale, 1997. *An Ecological Assessment of the United States Mid-Atlantic Region: A Landscape Atlas*, EPA/600/R-97/130, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C., 104 p.

Jones, K.B., A.C. Neale, M.S. Nash, R.D. Van Remortel, J.D. Wickham, K.H. Riitters, and R.V. O'Neill, 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: A multiple watershed study from the United States Mid-Atlantic Region, *Landscape Ecology*, 16:301–312.

Kennedy, M., and C.R. Meyers, 1977. *Spatial Information Systems: An Introduction*, Urban Studies Center, University of Louisville, Louisville, Kentucky, 97 p.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown, 1991. Development of a land-cover characteristics database for the conterminous U.S., *Photogrammetric Engineering & Remote Sensing*, 57(11):1453–1463.

Loveland, T.R., Z. Zhu, D.O. Ohlen, J.F. Brown, B.C. Reed, and L. Yang, 1999. An analysis of the IGBP global land-cover characterization process, *Photogrammetric Engineering & Remote Sensing*, 65(9):1021–1032.

Lowrance, R., L.S. Altier, J.D. Newbold, R.R. Schnabel, P.M. Groffman, J.M. Denver, D.L. Correll, J.W. Gilliam, J.L. Robinson, R.B. Brinsfield, K.W. Staver, W. Lucas, and A.H. Todd, 1997. Water quality of riparian forest buffers in Chesapeake Bay watersheds, *Environmental Management*, 21(5):687–712.

Lunetta, R.S., R.G. Congalton, L.K. Fenstermaker, J.R. Jensen, K.C. Mcgwire, and L.R. Tinney, 1991. Remote sensing and geographic information system data integration: Error sources and research issues, *Photogrammetric Engineering & Remote Sensing*, 57(6):677–687.

Lyon, L.J., 1979. Habitat effectiveness for elk as influenced by roads and cover, *Journal of Forestry*, 77:658–660.

———, 1983. Road density models describing habitat effectiveness for elk, *Journal of Forestry*, 81:592–595.

Mehaffey, M.H., M.S. Nash, T.G. Wade, C.M. Edmonds, D.W. Ebert, K.B. Jones, and A. Rager, 2001. *A Landscape Assessment of the Catskill/Delaware Watershed 1975–1998: New York City's Water Supply Watersheds*, EPA/600/R-01/075, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C., 117 p.

McGarigal, K., and B.J. Marks, 1995. *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure*, Gen. Tech. Rep. PNW-GTR-351, Pacific Northwest Research Station, USDA Forest Service, Portland, Oregon, 122 p.

O'Neill, R.V., C.G. Hunsaker, K.B. Jones, K.H. Riitters, J.D. Wickham, P. Schwarz, I.A. Goodman, B. Jackson, and W.S. Baillargeon, 1997. Monitoring environmental quality at the landscape scale, *Bioscience*, 47:513–520.

Osborne, L.L., and D.A. Kovacic, 1993. Riparian vegetated buffer strips in water-quality restoration and stream management, *Freshwater Biology*, 29:243–258.

Paul, J.F., R.L. Comeleo, and J. Copeland, 2002. Landscape metrics and estuarine sediment contaminants in the Mid-Atlantic and Southern New England regions, *Journal of Environmental Quality*, 31(3):836–845.

Peterjohn, W.T., and D.L. Correll, 1984. Nutrient dynamics in an agricultural watershed: Observations on the role of a riparian forest, *Ecology*, 65:1466–1475.

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

Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, S. Caicco, F. D'erchia, T.C. Edwards, J. Ulliman, and R.G. Wright, 1993. *GAP Analysis: A Geographic Approach to the Protection of Biological Diversity*, Wildlife Monographs No. 123, The Wildlife Society, Washington, D.C., 41 p.

Scott, J.M., and M.D. Jennings, 1998. Large-area mapping of biodiversity, *Annals of the Missouri Botanical Garden*, 85:34–47.

Swift, B.L., 1984. Status of riparian ecosystems in the United States, *Water Resources Bulletin*, 20:223–228.

Theobald, D.M., 2000. Reducing linear and perimeter measurement errors in raster-based data, *Cartography and Geographic Information Science*, 27(2):111–116.

Thiel, R.P., 1985. Relationship between road densities and wolf habitat suitability in Wisconsin, *American Midland Naturalist*, 113:404–407.

USGS, 1989. *Digital Line Graphs from 1:100,000-Scale Maps. Data Users Guide 2*, National Mapping Division, U.S. Geological Survey, Reston, Virginia, 88 p.

Veregin, H., 1989. A review of error models for vector to raster conversion, *The Operational Geographer*, 7(1):11–15.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering & Remote Sensing*, 67:650–662.

Walsh, S.J., D.R. Lightfoot, and D.R. Butler, 1987. Recognition and assessment of error in geographic information systems, *Photogrammetric Engineering & Remote Sensing*, 53(10):1423–1430.

Wehde, M., 1982. Grid cell size in relation to errors in maps and inventories produced by computerized map processing, *Photogrammetric Engineering & Remote Sensing*, 48(8):1289–1298.

Wickham, J.D., K.H. Riitters, R.V. O'Neill, K.B. Jones, and T.G. Wade, 1996. Landscape "contagion" in raster and vector environments, *International Journal of Geographical Information Science*, 10(7):891–899.

Wickham, J.D., K.B. Jones, K.H. Riitters, R.D. Tankersley, Jr., R.V. O'Neill, E.R. Smith, A.C. Neale, and D.J. Chaloud, 1999. An integrated environmental assessment of the U.S. Mid-Atlantic Region, *Environmental Management*, 24(4):553–560.

Wickham, J.D., R.V. O'Neill, K.H. Riitters, E.R. Smith, T.G. Wade, and K.B. Jones, 2002. Geographic targeting of increases in nutrient export due to future urbanization, *Ecological Applications*, 12(1):93–106.