16. Terrain Form Analysis of Geographic Information Systems

Kenneth L. Kvamme
Arizona State Museum and Department of Anthropology, University of Arizona, Tucson, AZ 85721, U.S.A.

16.1 Introduction

In a recent paper dealing with Geographic Information Systems (GIS) and archaeology, Harris and Lock (1990) emphasize that this technology might be the ideal means to link together the two approaches that archaeologists have employed in spatial investigations: (1) the visual and subjective appraisal of map-type information, and (2) quantitative spatial analysis. This chapter follows through on this idea. The GIS products that are produced in the following sections are maps in every respect and most clearly show, or at least suggest, a variety of patterns and relationships in the data. At the same time GIS are powerful data generators, capable of producing vast amounts of quantitative information that can be employed in subsequent spatial-statistical analyses.

In the following sections I utilize quantitative GIS-produced data as input to statistical pattern-seeking tests. This approach is combined, concurrently, with visual appraisal of patterns in a variety of cartographic products created via GIS. There is a two-fold focus to this endeavour. The first is the use of GIS as a means to assess whether an archaeological distribution is random with respect to background features of interest. The second examines terrain-type data that can be generated or derived by GIS, and particularly on a new terrain measure that offers a more comprehensive means of analysis of archaeological location. In all this I attempt to show some of the steps necessary in employing GIS as an analysis tool as well as the benefits of conducting analyses in a GIS setting.

16.2 A brief history of landscape analysis

Since at least the 1960's there has been great interest in analyzing archaeological distributions with respect to background features. Most often examined are environmental conditions in order to ascertain if tendencies exist in archaeological distributions for specific soil classes, elevation ranges, or water proximities, for example (e.g. Thomas & Bettinger 1976; Shermer & Tiffany 1985; Kellogg 1987). Distributional analyses of archaeological location also have been undertaken with respect to a variety of factors reflecting the social environment, such as proximities to prehistoric central places or roads (e.g. Plog 1971; Hodder & Orton 1976:227).

The principles underlying the methodologies employed for analyzing such archaeological tendencies are the same regardless of the nature of the background phenomena investigated and have undergone an interesting evolution in recent years. The earliest settlement studies, beginning with Willey's (1953) Prehistoric Settlement Patterns in the Viru Valley, relied heavily on subjective impressions of locational pattern gained through field experience or from the visual inspection of archaeological distribution maps. In such studies claims were made that sites exhibit locational tendencies along river bottomlands, for high terrace ridges, or for productive soils, for example (see Trigger 1968; Butzer 1971). This visual approach still remains in force today and is the primary means employed by archaeologists in regional investigations. Its importance as a major component in contemporary spatial analysis was emphasized in the introductory section. Even for quantitatively inclined researchers, the visual inspection of patterns in data plots is an absolute necessity in modern work.

Quantitative studies of archaeological distributional patterns with respect to background variables, usually environmental, increased during the early 1970's. Data were measured or observed at site locations in the field or on map sheets. The Southwestern Anthropological Research Group (SARG) perhaps best reflects this orientation toward quantitative measurement where participating archaeologists in the American Southwest were required to obtain at newly discovered sites such data as distance to nearest water, elevation, ground steepness, soil type, distance to nearest pueblo village, or distance to nearest kiva ceremonial site (Plog 1981). The use of these kinds of data initially were rather descriptive, with a focus on tendencies in site classes of interest. For example, prehistoric locational tendencies were described for paleindian groups in New Mexico by Judge (1973), for Middle Woodland settlements in Illinois by Roper (1979), and for prehistoric Shoshoneans by Williams et al. (1973) and Thomas and Bettinger (1976).

Statistical procedures for testing the significance of observed patterns also were introduced during the 1970's although they were not universally adopted. Plog and Hill (1971) advocated the chi-square goodness-of-fit test which, besides offering a means to assess the significance of site distributions, also provided a conceptual leap in regional studies (see also Davey 1971; Hodder & Orton 1976:225). Previously, locational tendencies and patterns exhibited by classes of archaeological sites would be argued solely on the basis of measurements made at those sites. For example, if 70 percent of settlements in a region of study were located on south-facing slopes, it would be claimed that these sites exhibited a tendency or preference for south-facing ground. The chi-square test forced a comparison of archaeological distributions against the nature of the background region at large in order to claim the existence of pattern. To put this in perspective, if 70 percent of the entire land area in the foregoing example possessed southern orientations then the observed tendency of the sites could be argued to be simply a reflection of the background environment. Conversely, if, say, only 20 percent of the study area reflected a southern exposure then the significance of the observed site pattern would be indicated. Shennan
KENNETH L. KVAMME

(1988) provides an excellent and current summary of this methodology.

One drawback of the chi-square-based approach is that the data must be categorical in nature. This test therefore is ideally suited for analyzing archaeological distributions with respect to nominal-level variables like soils, geological, or vegetation classes. Continuous data, such as distance to water, elevation, slope, or distance to nearest road, however, must first be categorized in order to employ this test. This amounts to the throwing away of information, since the scale of measurement is reduced, which can lead to less powerful inferences and which is undesirable for other reasons (e.g. one can no longer talk about mean tendencies, locational variability, or employ more powerful parametric tests).

One solution to this problem is to characterize the background environment on a continuous variable through a random sample of points taken from the study region at large. In other words, within the area of study the continuous variable of interest is measured at random locations. This body of measurements constitutes a sample approximation of the background environment that can then be compared against the same variable measured at a sample of archaeological site locations using such two-sample statistical tests as t-tests, Smirnov tests, or Mann-Whitney tests (Kvamme 1985; Smirnov tests, or Mann-Whitney tests (Kvamme & Tiffany 1985; Kellogg 1987).

A second approach, conceptually superior and statistically advantageous because the background distribution no longer is approximated by a sample, was first illustrated by Hodder and Orton (1976:226). This method involves obtaining the cumulative distribution of a continuous variable over the entire land area under study. To accomplish this they actually had to categorize their continuous variable, distance to nearest Roman road, into several classes. Then, superimposing a fine-mesh grid over their study area, southern England, and counting the number of grid units in each distance class, the proportion of the total study area within each class could be ascertained. A graph of the cumulative proportion distribution was obtained by plotting these values and interpolating the remainder of the graph. The cumulative distribution of their archaeological phenomenon of interest, a sample of coin locations, then could be statistically compared against this background constant through use of the one-sample Kolmogorov goodness-of-fit test (Hodder & Orton 1976:226; Orton:pers comm; see Kvamme, Chapter 10 of this volume for further details about this study).

Needless to say, the amount of effort required for this approach is excessive. Lafferty (1981) is probably the only researcher that further pursued this methodology with manual methods. Using a slightly different approach he divided a study region into 3,857 grid cells (each 200m square) and in each several continuous variables, such as elevation, slope, and distance to nearest water, were measured from map sheets. Simply by sorting the 3,857 measurements on any variable with a computer, the cumulative distribution of these variables could be obtained over the entire background environment. Measurements for the archaeological site samples were taken from those grid squares in which they were located and Kolmogorov tests were employed to analyze distributional tendencies.

It should be obvious that a raster GIS readily lends itself to this methodology (Wansleeben 1988; Kvamme 1989). An entire study region can be digitally encoded in a raster or grid cell data structure. In each grid cell the GIS can provide measurements for continuous (or categorical) variables systematically, grid cell-by-grid cell, over entire regions. Thus, it is quite easy to obtain the cumulative background population distribution for any variable, as well as measurements for archaeological sample locations, in order to assess whether the samples are unusual with respect to the spatial population via the Kolmogorov or other one-sample tests (see Kvamme 1990a for elaboration of these tests in GIS contexts).

It is emphasized that the GIS grid cell size, or resolution, controls the accuracy of the results. As cell size is reduced the proportion of the total encoded region at or below some value, x, on a continuous scale moves closer to the true value. Although greater accuracy is achieved, it must be realized that computer storage requirements increase geometrically with decreased cell size. Moreover, at some point the accuracy of the original map from which the data were obtained must be considered beyond which further cell size reductions become superfluous.

16.3 GIS-based analysis of the Marana Agricultural Complex

To provide illustration of GIS-based locational analysis a database is employed that characterizes a small region near Marana, Arizona, located in the Sonoran Desert region of the American Southwest. The area represented is a prehistoric Hohokam agricultural field complex that was in use during the Classic Period (thirteenth and fourteenth centuries A.D.). This particular field has been intensively studied (Fish et al. 1985; 1990) and contains an abundance of features related to Hohokam agricultural practices including terraces, check dams, and the ubiquitous rock pile feature. The latter are of particular interest in this study.

Rock piles are simply circular mounds or rounded heaps of fist-sized cobbles. Most rock piles are less than 1.5m in diameter and .75m in height. They are extremely common throughout the Hohokam territory and multiple lines of evidence strongly suggest they were employed for the growing of agave, an economically important plant that provided food and fibre (Fish et al. 1990).

The plant-growing environment is enhanced by the rock piles. Soils in this region have a high clay content that forces rainfall to run off rather than penetrate the surface. The rock piles represent a relatively porous surface that allows absorption of upslope run-off as well as direct rainfall. Also, the rocks act like a mulch, preserving the interior moisture by inhibiting evaporation through the blockage of capillary action and solar radiation (Fish et al. 1990). This effect has been established experimentally by Evenari et al. (1971:260) and the continued response of modern
plants to these enhanced microhabitats is indicated by a substantially larger root biomass in rock piles than in nearby control plots. Additionally, experimental evidence suggests that the rock piles offer a degree of protection to young plants from predation by rodents (Fish et al. 1985).

Locationally, these rock pile features generally are perceived to occur near ridge tops and on gentle slopes to capture surface run-off (Fish et al. 1990). Field inspection and distribution maps also give the impression of a tendency for north-facing aspects, perhaps to reduce the negative effects of direct solar radiation on evaporation and young plant growth.

16.3a The GIS Data Base

A 400 × 400m region is employed that surrounds the Marana agricultural field complex designated AZ:AA:12:470(ASM), illustrated in Fish et al. (1990:203). A random sample of 50 rock pile features from this region forms the basis of this investigation. Aerial photogrammetry was employed to produce an elevation contour map with a 61cm (2 feet) contour interval for the region. This contour map (Fig. 16.1a) was electronically digitized and spatial interpolation methods (Kvamme 1990b) were applied to estimate an elevation every four meters over the study area, yielding a 100 × 100m matrix, or a Digital Elevation Model (DEM) containing 10,000 values (Fig. 16.1b). This DEM, together with the 50 rock pile locations, provide the data necessary for the subsequent analyses.

16.3b Analysis of Elevation, Slope, Aspect

The impression of rock pile locations noted above, together with visual inspection of Fig. 16.1, suggests that they tend to occur somewhat downslope, but near the tops of the many ridges that traverse the study area. One way to assess this proposition is simply to examine the rock pile distribution with respect to elevation, since the major ridge tops tend to occur at higher altitudes. The cumulative background distribution was obtained by considering all 10,000 elevations in the data base. Elevations also were extracted at the 50 sample rock pile locations. Using these data the Kolmogorov test does indeed indicate a significant locational tendency (D = .21; p < .05) on the part of the rock pile sample for higher elevations (Fig. 16.3a), suggesting a locational orientation near ridge-tops.

One of the most powerful features of GIS is their ability to generate new data from existing information. Based on interrelationships between neighbouring elevations in the DEM, it is a simple matter for GIS to obtain slope (ground steepness) or aspect (direction of ground facing) information, two commonly derived data types. A number of algorithms exist for doing this (Kvamme 1990b). The one employed here fits a least-squares plane to each elevation and its eight neighbours (i.e. a local 3 × 3 matrix) and computes the maximum gradient, or slope, and direction of maximum slope, or aspect, on this plane. With 10,000 elevations in the DEM, 10,000 slope values and 10,000 aspect values are produced, each stored as a separate 'layer' in the GIS data base. Gray-scale techniques allow portrayal of key features in these layers. In the slope image (Fig. 16.2a) level ground is lightly shaded and steep ground is darkly shaded; in the aspect layer north-facing ground is lightly shaded while south-facing ground is darkly shaded (Fig. 16.2b).

From a theoretical standpoint the rock piles should tend to be located on mildly sloping ground in order that they may encounter accumulated surface run-off from up-slope regions, if current interpretations of these features are correct. The impression gained from field inspection of these features is that most of them tend to be located on sloping ground but this pattern, if it exists, is less than clear in Fig. 16.2a. Resorting to the quantitative GIS data, and the Kolmogorov test, the rock piles illustrate a significant tendency (D = .41; p < .01) for somewhat steeper slopes (Fig. 16.3b).
Figure 16.2: The derived GIS data layers used in the Marana locational analyses: a) slope, b) aspect, c) the ridge-drainage index. The 50 rock pile locations are shown in each.

As for aspect, the azimuths were first transformed by collapsing the west half of the compass over the east to yield a scale that reflects orientation on a north-south axis. On the new scale zero indicates north, 90 east or west, 180 due south, and values less than 90 degrees indicate a north-facing aspect. While the mapped rock pile distribution (Fig. 16.2b), cumulative graphs (Fig. 16.3c), and field impressions all suggest a tendency for more northerly-facing aspects, with 30 percent of the entire study area but 42 percent of the rock piles facing northward, for example, this orientation does not achieve significance by the Kolmogorov test ($D = .14; p > .05$).

16.3c Statistical properties of Elevation, Slope and Aspect

Elevation and its products, slope and aspect, are three of the most commonly employed terrain variables in archaeology and elsewhere. One of their most interesting properties is that they tend not to be closely correlated in regions of heterogeneous terrain. That is, elevation, slope, and aspect often can be regarded as independent sources of information that describe different characteristics of landform. This can be more clearly understood if one considers that level (or steep) ground can occur at any elevation; so too can south (or north) facing slopes. Similarly, south or north facing ground can be nearly level or quite steep. In other words, given knowledge of ground steepness or aspect it is not possible to predict elevation in regions of well-mixed terrain.

Table 16.1: Correlation matrix (Pearson's $r$) for the four terrain data types at the Marana agricultural site.

<table>
<thead>
<tr>
<th></th>
<th>Elevation</th>
<th>Slope</th>
<th>Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspect</td>
<td>0.03</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>Ridge-Drainage</td>
<td>0.13</td>
<td>0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>Index</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The foregoing can easily be demonstrated through computation of Pearson's correlation coefficient, $r$, between the three data sets (Table 16.1). That independence is the general situation clearly is indicated with the highest absolute correlation equal to only $r = .23$, or about $100r^2 = 5.3$ percent of the variance in common.

This characteristic of terrain data may represent one of the few places in nature where one gets something for nothing. Although slope and aspect are derived entirely from the matrix of elevation values this new information generally is independent, bearing little relationship with its parent.

16.4 A new landform measure: the Ridge-Drainage Index

One of the great potential advantages of GIS is that they facilitate the exploration and creation of new data...
Figure 16.3: Kolmogorov-Smirnov one-sample tests for the four variables: a) elevation, b) slope, c) aspect, d) the ridge-drainage index. With 50 rock piles a maximum deviation of the sample distribution of $d=.192$ is statistically significant at the .05 level.

types. With the speed of the computer, and the fact that whole landscapes are in digital form (i.e. the DEM), it is quite easy to rapidly explore new landform algorithms by applying them to the terrain data and mapping and evaluating the results. This potential has been explored for some time in other disciplines.
The ridge-drainage index is a continuous measure of landform shape that provides an indication of how ridge-like or drainage-like a particular location is. Like slope and aspect it too is computed within a 3 × 3 matrix, or moving window of elevations, obtained from a DEM. This index defines the central elevation as a vertex, and extends the contour of that elevation to the perimeter of the 3 × 3 window, returning the angle described by these perimeter points that encompasses the lower portion of the window (Fig. 16.4a). The concept behind this index is more clearly illustrated in Fig. 16.4b. Here, a contour map is shown with characteristic index values. Drainages, with narrow contour 'angles', yield low values; ridge flanks or plain-like regions yield intermediate values; and ridge crests, with wide contour angles, yield high values (Fig. 16.4b). The index yields values that range from zero (a pit) to 360 degrees (a peak); higher values are more ridge-like; lower values are more drainage-like; and values around 180 are terrain-neutral.

The ridge-drainage index can be applied to the 10,000 values of the Marana DEM to derive a ridge-drainage surface. The result is pleasing (Fig. 16.2c). High values of the index (lightly shaded) clearly correspond with ridge tops, low values (darkly shaded) with drainages, and intermediate values (medium gray) with the ridge flanks and plains areas (Fig. 16.2c; compare with Fig. 16.1).

16.4a Desirable index properties

Because the ridge-drainage index yields a continuous result it is possible to apply 'cut-points' to the measurement scale to further explore and define the nature of drainage or ridge systems in a region of study. This is illustrated with the Marana data in Fig. 16.5. By mapping all locations (grid cells) with an index value less than or equal to 90 degrees, those places that are most drainage-like are indicated (Fig. 16.5a). A cut-point of 135 (i.e. all values less than or equal to 135 degrees) defines well the Marana drainage network (Fig. 16.5b). A cut-point of 180, the scale mid-point (and also the mean and median in regions of mixed terrain, see below), neatly divides the study area into those regions more drainage-like (concave) versus those areas more ridge-like (convex) in form (Fig. 16.5c). This process can be continued. In Fig. 16.5d a cut-point of 225 shows in white the remaining ridge-crest system. In essence, this index illustrates that ridges are the mirror image of drainages (compare Figs. 16.5a,d) and that this characteristic of land morphology can be considered as a continuum.

The ridge-drainage index is nicely behaved in a statistical sense. In regions of heterogeneous terrain (ridges and drainages), a symmetric, bell-shaped distribution invariably results (Fig. 16.6), which tends to be somewhat leptokurtic (more peaked than a normal distribution). In these settings the index midpoint, 180, is also the mean, median, and mode, about which the distribution is symmetric. This result indicates that equal portions of mixed landscapes are drainage-like and ridge-like in character. The minor spikes at both extremes, 0 and 360, point to the relatively few pits and peaks that occur in landscapes (note that in a DEM context many of these pits and peaks can be artificial as a result of interpolation processes; unwanted pits and peaks should be removed prior to application of this algorithm through smoothing or other procedures).

Importantly, the ridge-drainage index tends to be uncorrelated with the elevation data from which it was derived, as well as independent of slope and aspect (Table 16.1). The highest absolute correlation between the index and any of the variables is only \( r = 0.13 \). While this may seem counterintuitive initially, since we tend to think of drainages being low and ridges being high in altitude suggesting a relationship with elevation, it is also the case that many minor drainages (with low
Figure 16.5: Definition of drainage and rim conditions through applications of cut-points on the ridge-drainage index. Cut-points of a) 90, b) 135, c) 180, and d) 225 degrees are shown.

index values) occur at the highest elevations while subtle divides between drainages (high index values) occur in any elevational context.

I believe that the ridge-drainage index represents a new and fundamental landform data type that carries different information about terrain, because of its independence of the other data types, elevation, slope, and aspect. Consequently, it should provide a potentially important perspective in archaeological and other landscape studies. Regional analyses frequently examine archaeological distributions with respect to high points or ridge crests. For example, it is often argued that hunter-gatherer sites were located near high points in order to watch for game, or that high points offered greater defensive potential (Jochim 1976). Similarly, proximity to drainages is a common focus of study (Shermer & Tiffany 1985; Kellogg 1987) and valley bottoms generally are regarded as offering greater shelter potential (Euler & Chandler 1978). The ridge-drainage index allows these landform contexts to be quantitatively identified and archaeological distributions can be spatially analyzed with respect to this index.

16.4b Analysis of the Ridge-Drainage Index at Marana

Returning to the Marana agricultural complex, the previous analysis of elevation suggested a rock pile locational pattern near ridge crest situations because of their tendency for higher elevation in the study region (Fig. 16.3a). This tendency seems to be apparent in the elevation maps (Fig. 16.1) and, when the rock piles are viewed against the ridge-drainage surface, the same conclusion seems to hold (Fig. 16.2c). Quantitative analysis supports this conclusion. The rock piles illustrate a significant tendency ($D = .22; p < .05$) for higher ridge-drainage index values indicating that these features tend to be placed at locations somewhat more ridge-like in character (Fig. 16.3d). While 50 percent of the study area is ridge-like in form (values greater than 180), 60 percent of the rock piles occur in this situation, for example (Fig. 16.3d).
Interestingly, very few of the rock piles are located in the most ridge-like contexts, with only 10 percent possessing index values greater than 211 (Fig. 16.3d). This makes sense because very high index values indicate ridge crests that tend to possess level ground (compare Figs. 16.2a,c). The rock pile slope analysis (Fig. 16.3b) indicated a marked preference for steeper ground and the elevation data suggested locational tendencies for the ridge flanks, somewhat below the ridge crests (Fig. 16.1). It seems to be the case that rock piles were placed on relatively steep surfaces to capture run-off, high on the flanks of the major ridges, but below the crests with, perhaps, a slight tendency for avoiding direct southern exposures.

The ridge-drainage index (Figs. 16.2c, 16.3d) further indicates that a very narrow range of surface forms was exploited: fully three-fourths of the rock piles possess ridge index values between 168 and 211 degrees, pointing to a preference for nearly planar (neither convex or concave) surfaces, but not necessarily level ones. What advantages were gained for agave agriculture in these contexts presently is unclear. Concave or drainage-like surfaces (low index values) would tend to channel water run-off, perhaps leading to erosion or silting problems. Greatly convex surfaces, on the other hand, might lead water away from rock piles. The need for further study clearly is indicated.

16.5 Conclusions

In the foregoing I have presented the case that GIS facilitates the analysis of archaeological location with respect to background environmental (or other) variables. Analysis is easier to do because all the data are in computer form making quantitative manipulation and cartographic production easier. Moreover, the rapidity with which results can be obtained offers a revolution in the way researchers can do their work. While previously it might have taken days to complete any one of the above analyses through manual means, it now takes only minutes, freeing the researcher from much drudgery and allowing more time for purely creative thought. Hopefully, the substance of our work will improve!

The importance of the dual modes of regional investigation, visual inspection and quantitative analysis, is emphasized. Visual inspection of archaeological distribution maps provides an important source of information and insight that can complement quantitative analysis. For example, analysis revealed a tendency for high elevation, but it was only through the examination of plots and pictures that the ridge-flank pattern could be ascertained. Visual inspection, by itself, was shown to have weaknesses, however. In one analysis (slope) the distribution map was unclear relative to the outcome; in another (ridge-drainage index) the narrow range of locational tendencies that analysis revealed was not apparent in the distribution map. Thus, the two approaches indeed are complementary; one should not be undertaken without the other.

That GIS allows new types of information to be generated was amply illustrated by the ridge-drainage index. This new index, together with elevation, slope, and aspect, provides a generally independent description of terrain form that I believe will allow more comprehensive analyses of archaeological distributions with respect to the landscape.

Finally, the empirical analyses of the Marana agricultural field complex give us a better understanding of what is going on in terms of the rock pile distributions. In my view a general weakness of current archaeological model building and theory formulation is that so much of it occurs in the arm chair! Good models and theories are based on evidence (something archaeological theoreticians often lack); GIS linked with spatial analysis techniques can help to give us better evidence upon which solid models can be built. Armed with the results of the foregoing and other analyses it is hoped that improved models and
understanding of prehistoric Hohokam agricultural practices will be achieved.

Acknowledgements

I wish to thank Paul and Suzy Fish of the Arizona State Museum, University of Arizona, for their kind encouragement and support in using their Marana data. The ridge-drainage index is available in the TERRAIN PAC and TERRA computer programs, for PC-compatibles, from the author.

References


