THE USE OF GIS IN THE SPATIAL ANALYSIS OF AN ARCHAEOLOGICAL CAVE SITE

HOLLEY MOYES

University at Buffalo, Department of Anthropology, Box 610005 MFAC 360, Buffalo, NY 14261-0005 USA
Current Address: 609 Castillo, Santa Barbara, CA 93101 USA

Although archaeologists traditionally have viewed geographic information systems (GIS) as a tool for the investigation of large regions, its flexibility allows it to be used in non-traditional settings such as caves. Using the example of Actun Tunichil Muknal, a Terminal Classic Maya ceremonial cave in western Belize, this study demonstrates the utility of GIS as a tool for data display, visualization, exploration, and generation. Clustering of artifacts was accomplished by combining GIS technology with a K-means clustering analysis, and basic GIS functions were used to evaluate distances of artifact clusters to morphological features of the cave. Results of these analyses provided new insights into ancient Maya ritual cave use that would have been difficult to achieve by standard methods of map preparation and examination.

Archaeological studies using geographic information systems (GIS) have most often been employed in regional analyses (Petrie et al. 1995). One explanation is that the original GIS was designed by government planners to solve regional planning problems on large tracts of land. However, the general spatial infrastructure of GIS is not scale dependent (Aldenderfer 1996), and the flexibility of the system allows it to be used in smaller geographic spaces. For this reason, it is hardly surprising that archaeology is witnessing a growing trend of intrasite (within site) use that includes innovations in three-dimensionally recorded excavation data (Craig 2000; Edwards 2001; Levy et al. 2001).

Even though archaeological excavations are conducted in three-dimensional space, archaeologists traditionally conceive of units as discrete, horizontal, stratigraphic levels. Divisions between levels may be arbitrarily assigned or may represent temporal or cultural changes. The convention has been to map units as multiple levels using two-dimensional top plans (plan views). Because a GIS does not represent a major cognitive departure from paper mapping traditions, it easily represents the world as static, two-dimensional (or two and one-half dimensional) surfaces. In a GIS, each stratigraphic layer within a unit is represented as a separate view or data layer. This custom is acceptable to the archaeologist who has traditionally conceived of stratigraphic levels as discrete units of analysis. Often in intrasite analyses, archaeologists are interested in artifacts and features. For this reason, object or vector data models are commonly chosen to represent these entities. Objects and features are classed and each class receives its own coverage or theme that may or may not be displayed within the view.

Unfortunately, this cognitive model is not always appropriate for archaeology in caves. Although archaeologists may conduct cave excavations, in many instances artifacts and features are located in tunnels that may wind back upon themselves, creating a situation in which one sits directly on top of another. To illustrate this problem, imagine a winding stairway. The steps are vertically separated but horizontally overlapping. Problems arise in creating a GIS when tunnels and chambers vertically overlap because, in these cases, units of analysis are not layers of strata but continuous surfaces. Additionally, caves may contain complex topology in which objects share subsets of each other’s volumes. For instance, along the same wall, artifacts may sit on overhanging shelves positioned directly on top of undercut niches containing objects of interest. These types of spaces are difficult to model because they are three-dimensional. The commercially available GIS software cannot yet account for these real world situations, since it is not designed to both display and conduct quantitative analyses in truly three-dimensional space.

This problem may be addressed by treating overlapping tunnels as separate views within a GIS, in effect making each tunnel or chamber equivalent to a discrete entity. Each area of the site matrix could be analyzed separately but displayed on a single coordinate system. However, it is not an ideal solution because the spatial relationships between the entities are not maintained, since this method forces the user to make arbitrary distinctions as to where to “cut” the space into layers. In order to facilitate quantitative analyses, it would be possible to produce map projections of overlapping areas on the same grid system, but doing so would adversely affect visualization.

Although both of these solutions could work fairly well, their biggest drawback is that they compromise spatial relationships that may be important to analyses.

For display and visualization, dimensionality may be achieved using TINs (triangulated irregular network) or DEMs (Digital Elevation Models), but these produce two and one-half dimensional graphics. This can be visualized by imagining the draping of a cloth over a geographic area. Additionally, numerous other software programs can successfully represent and display objects in three dimensions. Of particular note is the CaveTools extension for ArcView that is capable of representing underground tunnels as lines (Pratt 1998; Szukalski 2001). Likewise CAD (computer assisted design) representa-
tion of caves can be drawn. However, none of these approaches has the ability to quantify points or conduct basic GIS functions such as the creation buffers. Although display should not be underestimated as an important tool of scientific visualization for the archaeologist, geo-referencing and quantitative analyses of the distribution and spatial patterning of objects (such as artifacts or other features) are of primary interest. These analyses require the use of the quantitative analysis functions of a GIS.

For now, the decision on whether to create a GIS for analyses and display is heavily predicated on the topology of the space and goals of the research. Despite some of the problems that may be encountered, a GIS is still the most powerful tool available for geo-referencing objects and conducting two-dimensional spatial analyses. Caves or areas of caves that can be represented by horizontal planes are particularly good candidates for the creation of geographic information systems because they require minimal adaptations.

A case study of the Main Chamber of Actun Tunichil Muknal (ATM), or Cave of the Stone Sepulcher, illustrates these issues. The cave is an ancient Maya site located on the Roaring Creek River in western Belize near Teakettle village (Fig.1). It was discovered and named by geomorphologist Thomas Miller (Miller 1989), who produced a map of the 5 km system. The Western Belize Regional Cave Project, under the direction of Jaime Awe, conducted archaeological investigations there in 1996-1998. The Main Chamber of the cave functioned as a ritual space for the ancient Maya during the Terminal Classic Period (A.D. 830-950).

The chamber is a high-level passage that splits off from the tunnel system 500 m from the cave entrance. It provided an almost ideal spatial context for the implementation of a GIS because there were no overlapping tunnels and few problem areas containing artifacts. The area could be represented as a two-dimensional flat plane, which made it possible to create a GIS using commercially available software without modifications.

THE PROJECT

WHY GIS?

Although archaeological research has demonstrated that the ancient Maya used caves as ritual spaces (Brady 1989; MacLeod & Puleston 1978; Thompson 1959), the nature of the actual rituals is unknown. Archaeologists suspect that there may have been considerable variation in ritual practice both between and within caves. Because artifact deposition in ritual contexts is not expected to be haphazard or structurally amorphous, an examination of their distributional patterns should reflect ritual structure.

The goal of the project was to analyze the placement of artifacts within the cave’s interior using visualization and quantitative methods. The project entailed geo-referencing and tallying artifacts, examining their spatial distribution, and assessing artifact proximity to the morphological features of the cave. Comparing results generated by the study with ethnographic, iconographic, and ethnohistoric data should help to clarify the function and meaning of ancient Maya caves.

Achieving project goals required a means of data visualization, a high level of accuracy in mapping and analysis, and
a method by which to group the objects. GIS was instrumental in solving problems and eliminating obstacles that stood in the way of each goal. For instance, there were a number of problems with visualization. In the field, visual assessments of artifact distributions were prevented due to the limited range of our lights, large size of the chamber, and complex topology of the space. The long axis of the chamber is 183 m long and varies from 5-35 m wide, producing a total floor area of ~4450 m² (Awe et al. 1996; Moyes 2001; Moyes & Awe 1998, 1999). Large areas of breakdown, stalagmitic columns, and isolated boulders partition the space. Simple visual inspection was impossible, and the only mechanism by which to view the area was via a map, which led to further difficulties.

A total of 1408 artifact fragments were distributed throughout the chamber. The objects were piece-plotted at a scale of 1:60 to enhance detail and accuracy. The resulting paper map measured almost 4 m long. Viewing the document was difficult, but to reduce its size sacrificed legibility and introduced new distortions. By creating a GIS, it was possible to view the entire chamber on a single screen. A further advantage was that the GIS allowed the viewer to zoom into areas for close-up views with greater detail at an infinite number of scales. Representation of the same data at different display scales permits better visual inspection for patterning.

Proximity or associations between objects was important to understanding ritual activity. Therefore, a number of buffer areas surrounding cave features were needed for the analysis. These items included a diversity of sizes ranging between 10 cm - 1.5 m. Although it was possible to render these by hand, both the time required for such an undertaking and the compromise in expected accuracy due to line thickness and human error was prohibitive. The GIS provided a high level of accuracy for analyses and had the capability to create a variety of buffer zones surrounding specific morphological features within minutes.

Another problem was the development of a method for the quantitative analysis of the artifact assemblage. Over 99% of the artifacts were broken fragments, with some having been smashed into halves whereas others were broken into as many as 30 pieces, thus causing unequal weighting of the data. An in situ examination of the artifacts revealed that the fragments from single objects were deposited in close spatial proximity to one another. Often several objects were stacked together in piles or scattered in what appeared to be intentional groupings.

Clusters of artifacts offered a better unit of measurement than artifact fragments because they solved the problem of unequal weighting and were likely to represent specific isolated events. In many areas, such as small niches or rimstone dam pools, clusters were well bounded, but in large open spaces they were sometimes more difficult to define. This necessitated a formal cluster program to determine the cluster configurations mathematically. To create optimal, accurate, geo-referenced clusters, GIS technology was combined with statistical techniques.

**Mapping the Chamber**

The chamber was mapped and artifacts recorded during the 1996-1997 field seasons. Due to the difficulty in accessibility and wet, humid, conditions, no electronic mapping devices were employed and measurements were taken using tape and compass. To record artifacts, a system of 1-m grids was drawn over the base map, each grid assigned a number. Grid squares were located in the cave, and artifacts were piece-plotted on grid maps. These data were recorded on data sheets that were transferred to the base map.

**Creating a GIS**

The success of the study relied heavily on the design of the system. Because the primary goal of the project was to measure distances between objects, a vector model, representing features as points, lines, and polygons, was used. Points were used to represent artifacts and lines or polygons to represent morphological features of the cave. One of the most difficult tasks was to categorize features. The primary consideration was to devise a typology that reflected how the ancient Maya might have conceptualized interior cave space. A review of ethnohistoric and ethnographic documents, as well as archaeological reports, provided guidance for the creation of feature classes. Based on these data, as well as personal observations, features were divided into the following categories: 1) alcoves; 2) walls & walkways; 3) boulders; 4) breakdown; 5) niches; 6) pools; and 7) stalagmitic or stalacto-stalagmitic columns.

For purposes of the study, alcoves were defined as recessed or partially enclosed areas, accessible for human entry, opening onto a room, passageway, or tunnel. Alcoves could be at floor levels, sub-floor levels, or elevated. An alcove could not open onto another alcove, therefore it was partitioned on three sides. The category of “walls and walkways” was defined as the exterior boundaries of the chamber, or vertical structures that created interior partitioning. These might include large stalagmitic or stalacto-stalagmitic columns, areas of breakdown, or any other feature that delineates a passable route. Boulders were defined as detached rocks larger than 10 inches (~16 cm) in diameter and breakdown as “the debris accumulated from the process of collapse of the ceiling or walls of a cave” (Gary et al. 1972: 86, 112). Niches were described as very small alcoves that did not permit human entry. In the cave, they are usually recesses within walls or spaces under rock overhangs. Pools referred to gour pools and areas in the cave where standing water was observed during flooding. Stalagmitic or stalacto-stalagmitic columns are large stalagmites, groups of fused stalagmites, or those that formed a union of a stalagmite with its complementary stalactite (Gary et al. 1972).

Each category was represented by its own coverage. Walls and walkways were represented in the GIS by line coverages, and all others by polygons. Pools were represented using both lines and polygons. Maps were digitized using PC Arc/Info. To prepare for digitizing, paper maps were color coded by feature category. The paper map was divided into 6 segments to fit the...
between 4 and 6 tics were placed in each segment, corresponding to the previously developed 1-m grid system used for in situ recording. Once digitized, coverages were inspected, errors in digitizing were corrected, and polygon labels were examined for accuracy. Maps were then appended, edge-matched and transformed using Arc/Info. Transformation was accomplished by using grid coordinates from the original grid system. Once the coverages were edited, topology was built using Arc/Info and imported into ArcView 3.1 for manipulations of attribute tables, analyses, and data display. Maps of the cave were generated in order to view both global (Fig. 2) and local (Fig. 3) artifact deposition.

**CLUSTER ANALYSIS**

**K-MEANS CLUSTERING**

Clusters of artifacts were generated based on pure locational analysis using the K-means cluster program developed by Kintigh and Ammerman (1982). This method was employed because it is a simple, non-hierarchical program that can be applied to two-dimensional spatial coordinates of a set of points. Its application in this context was to determine whether specific artifact classes could be formed into a set of groups based on their pure spatial location. These groups, should they exhibit robust patterning, could then be related to specific morphological features in the cave using a GIS. In this research context, the approach was superior to point pattern methods such as nearest neighbor analysis.
Point pattern methods are generally concerned with the evaluation of the degree to which the individual members of a single artifact class have a tendency to be distributed randomly, homogeneously, or clumped, across a space with reference only to members of that class (Bailey & Gatrell 1995). These methods are powerful because they are based on the assumption that the spatial relationship of the members of that single class of artifacts vis-à-vis one another is intrinsically more important than the degree of spatial proximity of those artifacts to members of different artifact classes.

In contrast, pure locational clustering is not specifically concerned with a single artifact class, but instead the degree to which members of different artifact classes are found in close spatial proximity. The content of these clusters can then be evaluated to gain insights into past behaviors. This approach has the advantage of not weighing a priori any specific artifact class more than another. Instead, the method seeks to define “natural” groupings of objects across a space. While it is necessary to acknowledge that these methods often impose a structure on a data set, experimental studies have shown that K-means clustering generally provides excellent recovery of known data structure, especially when patterning is strong within the data (Aldenderfer & Blashfield 1984).

The number of clusters to be generated by the K-means program is determined by the user. The K-means algorithm allocates each point to one of a specified number of clusters and attempts to minimize the global goodness-of-fit measures by using an SSE (sum squared error). This measure is the distance from each point to the centroid of the cluster. Some programs allow the operator to view plot files of the SSE data to determine which number of clusters used produces the best goodness-of-fit configuration, but these programs can handle only small datasets. In order to handle the large ATM dataset, it was necessary to run the program in SPSS. Unfortunately, SPSS does not generate SSE plots, and although SSEs were numerically generated, they were produced by using a linear function, ill suited to the ATM spatial data.

New techniques were developed to determine the ideal number of clusters used for the K-means analysis. Although one option was to estimate the number based on perusal of the data, this was rejected for two reasons. First, it would have introduced bias to the data and defeated the purpose of numerical clustering, and second, not all of the points were well clustered and decisions on the number of clusters present in these areas would have been difficult, if not arbitrary. To resolve these issues, another quantitative method, LDEN (local density analysis) was enlisted.

**LOCAL DENSITY ANALYSIS**

The LDEN, or local density analysis, proposed by Johnson (1976) is a global measure designed to compute densities of artifact classes within a fixed radius of each point. The GIS program was used to generate x,y spatial coordinates for the 1408 artifact fragments and a LDEN was conducted on the data. The LDEN was iterated in .25 m increments, beginning at 25 cm and increasing to 3 m. The program was directed to produce a plot file of the results. The plot file that showed the highest local density coefficients of the spatial data occurred within the 25-cm radii. Using ArcView, a 25 cm buffer was created around each of the 1408 artifact points, and overlapping buffers were dissolved by the program, resulting in 252 polygons.

**TEST OF BEST FIT**

The K-means analysis was then initiated using the spatial data (x,y coordinates) and directed to generate 252 clusters. Before importing these data into ArcView for further analysis, the number of clusters was tested for best fit against higher and lower numbered configurations using the coefficient of variation (CV). The CV is defined as the ratio of the standard deviation to the mean:

$$CV = \frac{s}{X}$$

It is used to compare variables with unequal means by comparing the relative variability of a frequency distribution. Relatively less dispersed variables have lower coefficients of variation.

K-means clusters were created for 8 cluster configurations using the spatial coordinates. These were based on the original number of clusters generated by the ArcView buffers (252) and included numbers above and below that value. The cluster configuration numbers chosen were: 240, 250, 251, 252, 253, 254, 255, and 264. Seven numbered clusters from each configuration generated by the K-means were chosen at random for analysis. The CVs for the x,y coordinates for cluster numbers of each cluster configuration were added together and compared. The result showed that cluster configuration 252 had the lowest combined CV (.026554), demonstrating least variability in the data, therefore producing the best goodness-of-fit.

**CREATING A CLUSTER COVERAGE IN GIS**

Using the 252 K-means cluster configuration, a cluster attribute table was produced in ArcView. Each of the 1408 artifacts was assigned a cluster number. Numbers were highlighted and polygons were created using artifact points as nodes. This graphic was converted to a shapefile and imported into ArcInfo. Topology was built and the newly built coverage was re-introduced into ArcView.

The advantage of the new cluster coverage was that the clusters were smaller and more clearly defined than those generated by dissolving buffers around individual artifact points in the GIS program. This provided more accurate units of analysis and allowed for a better spatial resolution. Figure 4 illustrates shape and size differences between the two sets of clusters.

Using the 252 K-means clusters, buffers surrounding cave features were created using GIS. Results were generated for a variety of buffer sizes ranging from 10 cm to 1 m. As one would expect, spatial overlap of artifact clusters was present at...
all size levels. It was noted in the field that many artifacts were placed in pools, but also that a significant amount of the floor space was covered with intermittent standing water. If the pools category was considered separately, then the remaining categories accounted for 94% of the artifact clusters at the 25 cm buffer level. This demonstrates a reduction in category overlap at this level, suggesting that this buffer level is the best fit because it is the closest to 100%. Results were obtained using 25 cm buffers on the following feature classes: walls & walkways, stalagmitic/stalacto-stalagmitic columns, and boulders. Niches and alcoves were evaluated by determining clusters that intersected features. Because pools were represented by both lines and polygons, clusters were tallied manually on a presence/absence basis. The data are reported as a percentage of the total number of clusters found in association with each feature class: 1) pools (60%); 2) walls & walkways (28%); 3) boulders (23%); 4) stalagmitic/stalacto-stalagmitic columns (17%); 5) niches (12%); 6) alcoves (7%); and breakdown (7%).

**FINDINGS**

The data created by the generation of buffers was instructive because artifact depositional patterns that were not immediately obvious appeared in quantitative analyses. For instance the number of objects placed near or on boulders was unexpected. Ethnographic literature and iconography suggest that freestanding stones were used as altars or benches for the deities. Although rocks in caves are not morphologically similar to altars used in surface contexts, the data suggest that they serve as analogs to these features (Moyes & Awe 1998; Moyes 2001). Additionally, the quantitative data generated will aid in establishing methodology to evaluate variation in the ritual function of sites, one of the most pertinent questions facing Maya cave archaeologists. Quantifying objects and assessing their placement in relation to morphological features of the cave provide concrete units of analyses. These analytical units may be used to compare sites within and outside of the region. This study also reinforces the importance of visualization in assessing artifact assemblages. Viewing artifact clusters was instrumental in determining global spatial patterns, and areas of intense as well as sparse usage were easily identified. For instance, analysis revealed that 24% of the clusters were located in the central area. This area also contained the greatest variation in artifact classes as well as 40% of the total artifact assemblage. This ritual center of the chamber was, in fact, demarcated by the ancient Maya who placed what may be a
Three-Stone-Hearth in the middle of the area (Moyes 2000). The Three-Stone-Hearth is a salient concept in Maya cosmology representing the center of the universe and relating to the creation of the world (Friedel et al. 1993). This finds suggest that centrality was an important feature in the use pattern of the Main Chamber, was instrumental in spurring an investigation into the possibilities that creation and renewal rituals took place in the cave (Moyes 2001). Although it has been well recognized that rituals conducted in caves were often related to rain or water deities, this new finding suggests that, in this instance, water rituals may have referenced the flood event of the Maya creation myth.

Three global cluster patterns were identified as well: concentrated clusters, linear distributions, and isolated clusters located in peripheral areas. By comparing these configurations with ethnographic and ethnohistoric data, it became clear that linear distributions were most likely to represent ritual pathways (Moyes & Awe 1998, 1999), and isolated clusters most likely functioned as boundary markers.

Much of our knowledge of modern Maya spatial cognition comes from the work of Hanks (1984, 1990) who studied the Maya of Yucatan. He recognized that there is an important cognitive spatial component at the heart of all ceremonies performed by shamans. The model is a quincuncial configuration based on the four cardinal directions and a central area. Hanks (1990) described the ritual cognitive model as a centroid surrounded by a four-sided polygonal structure whose sides are created by joining the four intercardinal points.

This cognitive model has ancient roots. Evidence for its presence among the pre-Columbian Maya can be found in the Codex Madrid, in the layout of tombs at Río Azul (Adams & Robichaux 1992), and in site construction typified by the twin pyramid complexes at Tikal (Ashmore 1991). Ethnohistoric texts demonstrate the use of an elaborated quincuncial model in rituals of foundation intended to identify and sanctify community boundaries (García-Zambrano 1994). Note the similarity of the placement of linear scatters and boundary markers in the Main Chamber to the spatial model of the rituals of foundation illustrated by García-Zambrano (Fig. 5). This agreement suggests that interior cave space was ritually bounded and that establishing these boundaries may have been an important means of ritually defining a social universe within the cave (Moyes & Awe 1999; Moyes 2001).

CONCLUSION

Despite disadvantages that two-dimensional analyses may present when working with three-dimensional spaces, GIS is still the most powerful tool available for the analysis and display of archaeological data at every spatial scale. Although studies such as this could have been accomplished using paper maps, analyses such as the creation of buffers would not have been undertaken due to the time involved as well as the loss of accuracy and precision entailed in hand-drawing these entities. Although visualization would have been possible using other programs, no other system was capable creating geo-referenced data and conducting quantitative analyses. As this study illustrates, the strength of a GIS is demonstrated by its utility as a tool for visualization, data exploration, and data generation. There is no doubt that the GIS of the future will possess three-dimensional capabilities that will open up new lines of inquiry for spaces such as topologically complex caves.

ACKNOWLEDGMENTS

The Western Belize Regional Cave Project was funded by a grant from the Social Science and Humanities Research Counsel of Canada to Jaime Awe. I would particularly like to thank Dr. Awe for his support and encouragement. The permit for the project was provided by the Belize Department of Archaeology, and thanks go to Allen Moore, Brian Woodeye and John Morris.

Special thanks is extended to the 1997 staff of the WBRC, Sherry Gibbs, Cameron Griffith, Christophe Helmke, Mike Mirro, Caitlin O’Grady, Vanessa Owens, Rhanju Song, Jeff Ransom, and Kay Sunihara. Additionally, I am grateful to Keith Kintigh for his assistance in the early stages of this work, and to Ezra Zubrow and Nathan Craig for helpful comments. Thanks also to James Brady for his encouragement and assistance in the development of my work. Finally, I would like to thank Mark Aldenderfer for his comments and suggestions in the preparation of this paper. Although I have received insights from a number of people and valuable criticism from five anonymous reviewers for which I am grateful, the content of this paper is my responsibility.

This article is based upon work supported by National Science Foundation Grant No. DGE 9870668, “Integrative Graduate Education and Research Training in Geographic Information Science,” awarded to the University at Buffalo.

REFERENCES

Awe, J., Gibbs, S., & Griffith, C., 1996, Evidence for Late Classic elite ritual activity at Actun Tunichil Muknal, Belize, Paper presented at the 61st Annual meeting of the Society for American Archaeology, Nashville, TN.
Hanks, W., 1984, Sancitification, structure, and experience in a Yucatec ritual event: Journal of American Folklore, v. 97, no. 384, p. 131-166.
Hanks, W., 1990, Referential practice: Language and lived space among the Maya: Chicago, University of Chicago Press.
Johnson, I., 1976, Contribution méthodologique à l’étude de la répartition des vestiges dans des niveaux archéologiques [Thesis for obtaining a Diplome des Études Superieurs]: Université de Bordeaux I.