Settlement Pattern Modelling through Boolean Overlays of Social and Environmental Variables

Zoran Stancič
Centre for Scientific Research of the Slovenian Academy of Sciences and Arts.
Gosposka 13, 1000 Ljubljana, Slovenia.
E-mail: zoran@zrc-sazu.si

Kenneth L. Kvamme
Department of Archaeology, Boston University.
675 Commonwealth Avenue, Boston, MA 02215, USA.
E-mail: kvamme@crsa.bu.edu

Abstract

Robust, multivariate statistical methods frequently cannot be applied to model archaeological settlement distributions, in contexts, where the number of sites is small. One alternative is to apply simple Boolean logic, to combine variables, that have been shown to have a bearing on settlement locations. This paper focuses on the modelling of archaeological settlement distributions, through the use of simple Boolean overlays in GIS, a method, that is, by no means, new. The sites investigated are a sample of Bronze Age hillforts, from the island of Brač, in central Dalmatia, Croatia. What is new in this paper is that particular attention is paid to social variables, as "predictors" of settlement location, a domain, too frequently overlooked in modelling studies. A number of social variables are investigated, and their importance is statistically tested. At the same time, an environmental perspective is not sacrificed, for multiple environmental correlates, of settlement location, are shown to exist, as well. Consequently, this paper recognizes the importance of social and environmental realms, to human location behaviour, and shows that models of high predictive power can be achieved, when variables from both domains are simultaneously considered.

Introduction and background

Predictive models have a fairly long tradition in American archaeology, where they have been employed for cultural resource management and planning purposes. Given the vast tracts of federally administered lands in the United States, predictive models have been used to generalize the likelihood of site locations, over wide areas, on the basis of limited archaeological surface surveys (Warren, 1990; Judge and Sebastian, 1988; Kvamme, 1992). In European countries, archaeological legislation is generally different, and allows archaeologists to do their work, easily, on privately owned land. Nevertheless, the application of predictive modelling in Europe has been very limited, and often, simplistic. One objective of this paper is to discuss conceptual issues, in archaeological predictive modelling. A second objective is to examine how social variables can be incorporated into these models. Applications in the USA have been criticized, because, for the most part, only natural environmental data have been employed, in archaeological spatial analysis (Wheatley, 1995). We, therefore, suggest several ways of quantifying social variables, and of using them in archaeological, predictive models.

After a general description of data, from the island of Brač, in central Dalmatia, Croatia, and an examination of data quality issues, a simple model for hillfort locations is presented, based on Boolean methods and map algebra. In a final section, the modelling technique is assessed and general conclusions are drawn.

Although it is not within the scope of this paper, to fully discuss the theory of predictive modelling, in general, two approaches have been defined: inductive and deductive. In an inductively based model one begins with basic archaeological data and attempts to build a model, based on patterns in the database. In a deductive approach, one starts with theoretical knowledge and an understanding of archaeological cultures, on the synthetic level, and tries to deduce conclusions, about past settlement and land use logic (further details can be found in Dalla Bona, 1994). To implement models, one can apply the GIS technique of Boolean overlays of relevant variables, map algebra methods, or utilize multiple regression methods. The last approach, usually associated with an inductive modelling stance, is very powerful and can give insight into relationships, between the individual variables analyzed. Regression analyses, however, require large site samples. Unfortunately, large samples are not the usual case, in many studies, where knowledge of land use or settlements is limited to several locations only.

The study area and database

The Island of Brač

The data set, examined here, is from the island of Brač, in central Dalmatia, Croatia. The central Dalmatian Islands have been the subject of extensive field survey, for more than a decade. The Adriatic Islands Project (Gaffney, et al., 1997) has conducted field surveys and excavations on several islands, ranging from the small island of Palagruža, in the center of the Adriatic Sea, through the islands of Šolta, Vis, Hvar, and up to the largest island in the region, Brač. Modelling techniques were tested on Brač. A fairly large number of Bronze Age hillforts exist on the part of the island
examined.

It is important to stress that islands are ideal for archaeological, spatial analysis, since they are well-defined spatial entities, where land territories are easily conceptualized.

Brać is the largest of the central Dalmatian Islands, with a total surface area of 395 km$^2$. It is of elliptical shape, with its long axis, oriented east-west, and measuring around 36 km in length, while the shorter axis is about 12 km long (Figure 1). All of central Dalmatia is characterized by rather dramatic relief, and Brać offers no exception. The highest peak on Brać is Vidova gora, measuring 778 m, above sea level; it is also the highest peak in all the Adriatic Islands. On the other hand, the geology of the island is fairly monotonous, being comprised of Cretaceous limestone and dolomite. Soft Eocene deposits can be found, only in small areas, on the southern coast, while quaternary deposits can be found in most valleys and in numerous karst dolinas. The best soils were developed on these two geological bases.

The climate of the island is typically Mediterranean, with mild winters and hot summers. Yet, despite its small size, there are variations in microclimate. First, because of northern winds, the north coast is a bit colder. Although the average summer temperature is 16°C, the temperature drops about 0.6°C, for every 100 meters of elevation rise. The relief has a similar impact on precipitation, which is nearly all rain. While on the western tip of the island, the average rainfall is 799 mm, per square meter, yearly, in Pracnice. the highest in all the Adriatic Islands. On the other hand, several hillforts seem to be rather isolated, with no substantial island surface has been eliminated. This study area encloses what is considered to be, the center of Bronze Age activities, within which, 107 barrows and eight hillforts are located.

Archaeological Data Quality

One question, that must be addressed, is the contemporaneity of the sites we are analyzing. The data used in this research were obtained through archaeological field survey. On Brac, only a single hillfort has been excavated, with results that remain unpublished, and provide limited insight into the Iron Age, only (Marović and Nikolanci, 1977). On the basis of comparable data, from Bronze Age sites, on the neighboring islands and mainland, all eight hillforts were identified as major Bronze Age settlements. Only extensive excavations could provide detailed chronologies of each individual site.

Similar problems occur when barrows are considered. It is generally agreed that most of the barrows, in the central Adriatic, are dated to the Bronze Age. Yet, we cannot assume they were constructed, in a short period of time. If we consider the labour needed to build a 20 meter diameter barrow three meters in height, it is clear that they must have been gradually built, over long periods of time. There is also the problem of the function of barrows in the Bronze Age. Several functions have been assigned to them, from obvious ones like burials, to functions as landmarks, or ritual places (Gaffney and Stančić, 1991).

Some differentiation of hillforts, stemming from their function, is suggested in the distribution map of hillforts and barrows (Figure 2). It is clear that some hillfort sites are virtually surrounded, by the more numerous barrows. On the other hand, several hillforts seem to be rather isolated, with only several barrows in association with them. As an
analytical start-point, and since some of these sites must have coexisted, it was decided to treat all barrows and hillforts as contemporaneous. Moreover, each respective group was treated as a functionally homogeneous unit. The study area included eight Bronze Age hillforts and 107 barrows. While the number of barrows is fairly large, providing a good database for predictive modelling, the eight hillforts are more problematic. For the latter, the focus of this paper, GIS-based Boolean intersection and map algebra methods are employed, as the model-building mechanisms. The advantage of these approaches, over complex multivariate statistical techniques, is that they are easy to understand, to implement, and they can be applied to very small samples.

During the analysis, another problem was encountered, stemming from variable site sizes and the accuracy of their recording. All the sites, under consideration, are fairly large. The smallest barrows have surfaces of more than 20 m² and can be up to 400 m² in area. In the raster GIS, with a cell size of 30 x 30 m, the size of these sites can be neglected, and they can be easily treated as points. This is not the case with the hillforts, however. Hillforts can be extremely large, covering several hectares. Hillfort locations were recorded, using only a single spatial coordinate, however. Unfortunately, this coordinate was not always the center, or the highest point, of the site. During the analysis, this caused several problems. For example, when the slope of the hillfort sites was analyzed, it was found that one site was located on very steep terrain, with the slope measuring over 60 percent. Evidently, the location of this hillfort must have been recorded, on its southern ramparts, which were at the edge of a dramatic slope. Despite these problems, it was decided to leave the data unchanged.

**Natural Environmental Data**

The most important natural environmental data source, in GIS-based spatial research, is usually the digital elevation model (DEM) and its derivatives. A DEM was created, using contour lines from photogrammetrically produced maps, at a scale of 1:25,000. Contour lines were digitized and used for the interpolation of the DEM. A number of terrain data types were derived from the DEM and used in the analysis. These included slope, a local relief measure, as well as variables, more related to the social domain, like the size of hillfort territories and the intervisibility between them. It was decided to also consider other natural, environmental data like soil quality, which must have influenced land use patterns, in the past.

The natural, environmental data used in the analyses represents the present environment. A fundamental question is how well the environmental data represents the period analyzed. Relief, of course, has not changed substantially, since the end of the last glaciation, but this can not be said of all the environmental information. Water springs and natural ponds were recorded during the field survey, but these were felt to have very limited potential, in our analysis. It appeared that some hillforts existed without any water resources, for example. It was, therefore, obvious that either a substantial number of water resources remained unrecorded, or that they have changed. Such a change may be due to dramatic differences in vegetation cover, from the prehistoric condition.

Events at the end of the last century provide considerable insight into this phenomena. At this time, the central Dalmatian Islands were intensively used for viticulture and agriculture, due to an increased demand, caused by a vine disease in the western Mediterranean. Previously unused land, covered with shrubs and grass, was cleared, terraced, and changed into vineyards, dramatically reducing vegetation cover. During fall storms, large quantities of rain caused creeks to flood, and several people were actually killed. Today, with Brac covered in dense Mediterranean shrubs, it is hard to believe that a dry valley, with no water throughout the year, grew to a substantial stream that flooded and took human lives!

Finally, the distribution of soils must have played a crucial role, in the establishment of Bronze Age settlement patterns. Yet, the soils database, produced from very limited fieldwork, was not as good as the one, for the neighboring island of Hvar (Gaffney and Stančić, 1991). At a scale of 1:200,000, the soils map was barely usable, for any kind of analysis. It was, therefore, decided to derive soils information from satellite imagery. A Thematic Mapper image was used to produce soils data, which, despite the mixed signal from the vegetation, proved to be of much better quality than the original map. The question of temporal changes in soils cover remains important, however. Comparative analyses of soils on the mainland (Shiel and Chapman 1988) show significant changes, since the Bronze Age, suggesting that the soil data should be approached with extreme caution.

**Boolean models, map algebra, and hillforts**

**Introduction**

The Bronze Age hillforts seemed a perfect example of sites, against which the impact of variables, that might influence their location, could be evaluated. These sites are interpreted to be important settlements, which were located on hilltops, within some kind of defensive structures. They probably represent the highest level of a hierarchy of settlements in the region (Marović, 1981). Settlements with a smaller number of huts, or even groups of huts, must have existed at other locations. By definition, hillforts must be located on hilltops, and the summit must have a level slope, to be adequate for the settlement. A number of social factors must also have influenced their locations. The most obvious is a distance factor, to the nearest hillfort. It is generally agreed that hillforts were associated with territories, within which, there should not be another hillfort. The main objective of this study is to examine and model the relationship, between hillfort sites and their natural and social environment.

The basic logic, behind a predictive model, is rather simple. First, a set of variables, which are considered to have influenced site locations, in this case hillforts, are defined. Then, hillfort characteristics, on each variable, are compared against locations, without hillforts. On the basis of this comparison, a threshold value is defined for each variable. This threshold is used to create a binary layer, for each variable, that indicates hillfort-like versus hillfort-unlike locations. Finally, the binary layers are combined, through a Boolean intersection, or a map algebra summation. The result is a simple model for hillfort locations, and such a model, if
it fits the known hillforts well, has a predictive capacity, in that it may be employed as a prospecting tool for new hillforts. Although effective, if the proper variables and thresholds are isolated, this approach has some obvious disadvantages, compared to a multivariate, statistical approach. The most obvious one is that each variable is treated separately, so at the end, we cannot quantify the relative influence of each variable. It would be intriguing, for example, to compare the weights of the natural variables, against the social ones. Such an approach would allow discussion of cultural versus natural environmental factors, influencing settlement choice (Kvamme, 1997).

A model is always a simplification of the real world. Consequently, some variables may be omitted, because we simply are not aware of their influence. Other variables are not employed in a model, because they are difficult, or impossible, to measure. Based on our knowledge of prehistoric hillforts in the region, and our experience obtained during field survey, a list of relevant variables, influencing hillfort locations, was made (Table 1).

**Social Variables**

Four variables fall within a group of socially related variables, influencing hillfort locations. The first is the distance between hillforts. It is clear, in the hillfort spatial distribution (Figure 2), that they are regularly distributed, probably because of their need for economic support territories. These territories require a minimum distance between hillforts. It was calculated that the minimum distance, between two hillforts, is about 1600 meters, meaning that each hillfort should have at least an 800 meter, exclusive buffer zone.

A second, important social variable was the intervisibility between hillforts. It is quite likely that hillforts were positioned at locations, that would offer visual control over large areas, that included their own territories, plus other hillforts. It was, therefore, hypothesized that hillforts were located, where it was possible to see many other hillforts. To test this possibility, a viewshed area from each hillfort was calculated, and these areas were summed, to create a cumulative viewshed (Wheatley, 1995). The cumulative viewshed is a thematic representation of hillfort intervisibility, where higher values indicate that more hillforts are visible. Our eight hillforts appear to be located in places, highly visible from other hillforts. While the average was 1.2 visible hillforts, for all locations within the study area, an average of 3.9 hillforts were visible, from the eight hillfort locations.

The distance from the sea was a third, socially related variable, examined. Bronze Age hillforts were uniformly located a considerable distance from the coast, probably because larger distances meant greater safety from pirates and raiders (hence, the underlying causal mechanism, behind this variable, is a social phenomenon). The distance from the coast was measured using linear and slope modified distances. In both cases, a relationship between the distance, from the coast and hillforts, is shown. The average distance from the coast, considering all locations on the island, is around 2300 meters. Yet, the hillforts appear to be located far from the coast, with an average distance of about 3200 meters. A similar pattern is found, using a cost surface approach, calibrated to walking times. The average distance to the coast, for the whole study area, is around a three hour walk, compared to an average of 4.5 hours to the hillforts. It was decided that cost surface approach results, best represent real world circumstances.

The final, socially related variable is based on the location of barrows. During the field work, it was noticed that barrows, typically appear at certain regular distances from the hillforts, but not necessarily proximate to the hillforts. Consequently, a gravity model was constructed, through an overlay of cost surface distances, computed from each of the more than 100 barrows in the study area. When the hillforts were compared against this surface, it was found that they typically occur, within a limited range of distances from the barrows, but with some minimum distance from them, maintained.

**Environmental Variables**

The natural, environmental data considered for hillfort locational modelling, include slope, a ridge-drainage index, a rim index, and a relief below measure. Slope is simply, the gradient or ground steepness of the terrain, at a hillfort. It was assumed that hillforts are located on hilltops, with fairly level ground, to accommodate habitation and day-to-day activities. However, the average slope of the study area is about 17 percent, while the mean slope of the hillforts is 20 percent. As mentioned earlier, this result is due to the fact, that the spatial coordinates of some hillforts were recorded on the sloping ramparts, and not in the center of the hillforts.

The remaining variables attempt to quantify the principal characteristics of hillfort location: a dominant elevation above a flat terrain. The ridge-drainage index actually calculates a "viewing angle", of a location. At drainage-like locations, the viewing angle tends to be much smaller than 180 degrees; on ridge locations, it is much larger; on a peak, it is close to 360 degrees. The rim index determines the volume, within a specified distance from a central point above the ground surface, and below a plane, an arbitrary 100 m above the point. If the point is on a peak, or at a rim above a steep drop, it tends to yield a higher index than that obtained in a valley-like context. Finally, relief below quantifies the elevation range of the surrounding area. It is simply the maximum elevation drop, within 300 m of a locus. These variable points are described in more detail, elsewhere (Kvamme, 1992); we feel their combination gives a good representation of a hillfort, within its natural environmental context.

We also wanted to examine the relationship, between soils and the hillfort locations. The availability of adequate soils, for agriculture, is an important limiting factor, in the general distribution of settlements in the central Adriatic. Good quality soils are usually limited to karst dolinas, alluvial valleys, and some minor areas of Eocene geology. Because good soils maps were not available, our research utilized a classified Thematic Mapper image, from July, 1993. Through this image, several classes of soils were easily interpreted. The good quality soils, which are intensively used today for agriculture, were easily identified. So, too, were areas of very poor soils, which are mostly abandoned.
and have very scarce vegetation. The intermediate soils were rather difficult to interpret, however. It often happens that a dense Mediterranean shrub has overgrown the area, and it is rather problematic to separate soils, with the same vegetation cover.

Our general hypothesis was, that each hillfort should contain some good quality soils, within its catchment. Catchments were defined as circles of 800 m radius, around each hillfort. With only eight sites and a weak correlation, between hillfort territories and soils, it was decided to employ Monte Carlo methods, to compare the hillfort territory soil quality, against 99 randomly generated, sample territories, each of eight locations. A weak, but significant, correlation (at the five percent level of significance) was found. Nevertheless, due to the weak pattern, which might stem from the poor quality of the soils data, it was decided to drop the soils data, for modelling purposes.

Modelling Thresholds and Performances

On the basis of the foregoing investigations, a threshold value was defined on each variable, that met certain criteria of relevance to hillfort locations. These threshold values allowed generation of a binary information layer, for each variable, where all locations that met the threshold were assigned, a value of one; zeros were assigned otherwise (Figure 3). These threshold values could be defined, at any point on a variable's measurement scale, but we employed a liberal criterion, such that all the known site locations would be captured by the threshold.

During the creation of these binary layers, the areas assigned a value of zero were closely monitored, as an "index" of performance. The larger this area, the more a variable contributes to eliminating locations, unlike the hillforts. In Table 2, the threshold values, as well as the performance of each predictor, is indicated. It is clear that the strongest predictor of hillforts are the rim index and the ridge-drainage index. With these variables as hillfort models, alone, we loose 80% or more of the landscape! Of the natural, environmental variables, slope performs the worst, probably due to the inaccuracy of the data, discussed earlier. With a threshold steepness of 60%, a reduction of only 3% of the study area is achieved. Many of the social variables are rather strong predictors, as well (Table 2). Each of them enables us to reduce the area, of possible site locations, by some 50%, which we consider to be very good performance.

Building the Model

A predictive model for hillfort locations was constructed simply, by summing the binary layers, using GIS-based, map algebra methods. The resulting model layer, therefore, must range from zero to eight (the former occurs, if none of the conditions are met; while the latter arises, when all conditions are met). The resulting model is presented in Figure 4. We note that locations, that meet all eight thresholds, represent the Boolean intersection of the eight binary layers. Moreover, the summing operation offers the advantage that each location, within the study area, is rated incrementally: locations assigned a value of "8" possess more hillfort-like characteristics, than a location assigned a value of "7", and so on. The ranking, of course, represents the number of thresholds met, so a location that meets six, seven, or eight of them, can be considered to be good possibilities for hillfort locations (see Williams, et al., 1973, for a similar methodology).

In the present case, all the known hillforts are within the areas of highest ranking, or likelihood of containing hillforts (i.e., a rank of 8). This area includes only 0.22 percent of the entire study region! Yet, these results can be improved, even further, because some of the locations of highest ranking are very small in size. Small areas, of less than 1000 m², do not provide enough space to contain a hillfort. Consequently, if a size restriction is also applied, the number of possible new hillforts, indicated by the model's highest ranking, is reduced to only a handful, which can be easily examined in the field.

Summary and conclusions

Boolean and map algebra methods, for producing predictive models, are straightforward, easy to implement and to understand, compared to some multivariate, statistical procedures. On the basis of threshold values, derived from analysis, binary layers for each variable are generated. A predictive model is essentially a combination of these binary layers. The overall performance of each variable can be assessed, by the amount of the study area it helps to eliminate, as unlikely loci for possible site locations. These methods, however, lack insight into correlations, associations, and the relative importance of the variables, that statistically based procedures can provide. On the other hand, these procedures can be used in contexts, with a small number of sites. Due to their simplicity and good performance, even on smaller data sets, Boolean and map algebra methods represent effective tools for creating predictive models.

Given that archaeological sites were placed, according to a variety of natural environmental and social constraints, predictive models can provide theoretical insights and understandings into past systems of land use and occupation, by making associations, with these features and the overall settlement patterns clear. Moreover, model results can be used in more innovative ways, by allowing tests of various hypotheses. With the widespread use of GIS, increasing computerization of archaeological sites and monuments records, and a growing demand for archaeological sensitivity maps, one can expect that predictive modelling will be a fast developing field, of quantitative and computer archaeology.

Acknowledgements

The Slovenian Science Foundation granted a scholarship to Zoran Stančič, for the research behind this paper, at the Department of Archaeology, at Boston University. The European Commission, DGXII, financed the field survey on the island of Brač, Croatia. Computerization of the site and monument records were partially supported, through the Research Support Scheme of the OSI/HESP, grant No. 1363/1997. Special thanks go to the Ministry of Science and Technology of the Republic of Slovenia, for supporting bilateral cooperation with the Republic of Croatia.
References cited


Tables

<table>
<thead>
<tr>
<th>variable</th>
<th>background mean</th>
<th>background s.d.</th>
<th>location mean</th>
<th>location s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum distance between hillforts</td>
<td>/</td>
<td>/</td>
<td>1600</td>
<td>/</td>
</tr>
<tr>
<td>intervisibility</td>
<td>1.2</td>
<td>1.6</td>
<td>3.9</td>
<td>1.7</td>
</tr>
<tr>
<td>distance from the coast</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a.linear distance</td>
<td>2300m</td>
<td>1500</td>
<td>3200m</td>
<td>1300</td>
</tr>
<tr>
<td>b.cost surface</td>
<td>190 min</td>
<td>110</td>
<td>270min</td>
<td>50</td>
</tr>
<tr>
<td>barrows gravity model</td>
<td>28759</td>
<td>10081</td>
<td>22131</td>
<td>5602</td>
</tr>
<tr>
<td>slope</td>
<td>17%</td>
<td>16</td>
<td>20%</td>
<td>18</td>
</tr>
<tr>
<td>ridge/drainage index</td>
<td>181</td>
<td>28</td>
<td>257</td>
<td>57</td>
</tr>
<tr>
<td>rim index</td>
<td>870</td>
<td>80</td>
<td>1065</td>
<td>97</td>
</tr>
<tr>
<td>relief bellow index</td>
<td>27</td>
<td>22</td>
<td>43</td>
<td>25</td>
</tr>
</tbody>
</table>
Table 2. Variables used in hillfort predictive model with the threshold values and performance.

<table>
<thead>
<tr>
<th>variable</th>
<th>threshold value</th>
<th>performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum distance between</td>
<td>$x \geq 1600m$</td>
<td>47%</td>
</tr>
<tr>
<td>intervisibility</td>
<td>$x \geq 1$</td>
<td>52%</td>
</tr>
<tr>
<td>cost surface distance from the coast</td>
<td>$x \geq 220\text{min}$</td>
<td>57%</td>
</tr>
<tr>
<td>barrows gravity model</td>
<td>$16000 \leq x \leq 32000$</td>
<td>43%</td>
</tr>
<tr>
<td>slope</td>
<td>$x \leq 60%$</td>
<td>3%</td>
</tr>
<tr>
<td>ridge/drainage index</td>
<td>$x \geq 198$</td>
<td>80%</td>
</tr>
<tr>
<td>rim index</td>
<td>$x \geq 950$</td>
<td>86%</td>
</tr>
<tr>
<td>relief bellow index</td>
<td>$x \geq 14$</td>
<td>30%</td>
</tr>
</tbody>
</table>

List of Figures in CD-ROM.

**Figure 1.** The island of Brač.
**Figure 2.** The general distribution of hillfords and barrows in the working area.
**Figure 3.** Individual binary layers, based on defined threshold points, of the eight social and environmental variables.
**Figure 4.** Predictive model for hillfort locations, as a Boolean overlay of eight social and natural, environment variables.