Original article

Locating Macedonian tombs using predictive modelling

Aikaterini Balla a, *, Gerasimos Pavlogeorgatos a, Despoina Tsiafakis b, 1, George Pavlidis b, 1

a Department of Cultural Technology and Communication, University of the Aegean, University Hill, building of Geography, off. 2.14, 81-100 Mytilene, Lesvos, Greece
b Cultural and Educational Technology Institute, ATHENA Research Centre, 58 Tsimiski Str., 67100 Xanthi, Greece

A R T I C L E I N F O

Article history:
Received 28 November 2011
Accepted 15 October 2012
Available online xxx

Keywords:
Archaeology
Predictive modelling
GIS
Fuzzy logic
Macedonian tombs

A B S T R A C T

Northern Greece is the main region where more than 90 Macedonian tombs, dated in the late classical and Hellenistic period, have been discovered. Geospatial data processing technologies (GIS), predictive modelling techniques and fuzzy logic were applied to the study area in order to create a predictive model that would be able to provide map regions assigned with specified probability of Macedonian tombs’ occurrence. The model was tested extensively and was validated using a commonly used predictive gain. The results and the evaluation of the model proved the efficiency of its predictive ability in providing answers to a series of questions related to the problem at hand (archaeological research, cultural resource management and protection, land use, etc.).

© 2012 Published by Elsevier Masson SAS.

1. Research aims

Macedonian tombs constitute a special category of burial monuments built mainly in Northern Greece during late classical and Hellenistic period (middle of 4th century B.C. – middle of 2nd century B.C.). This paper presents an attempt to examine their common spatial attributes and their distribution in the study area, in order to create a predictive model that would identify areas of possible Macedonian tombs’ occurrence. Such model would enrich archaeological knowledge about ancient Macedonian culture and would contribute to the study of ancient topography, as the discovery of new tombs can result in finding yet undiscovered settlements, roads etc. related to the burial monuments. Additionally, the knowledge of areas expected to contain Macedonian tombs can be exploited in development planning, in order to prevent possible future damage of those monuments by modern development interventions. Furthermore, the proposed model can be used as an efficient solution to the lack of funding by minimizing the number of trial excavations and by indicating specific areas that are of high probability to result in finding undiscovered archaeological remains.

2. Introduction

The application of GIS technologies in archaeology over the past recent years has yielded important expertise that can be successfully exploited by both archaeological research and Cultural Resource Management (CRM). The application that set GIS as a mainstream tool in the field of archaeology is predictive modelling. The basic principle upon which this scientific field was based on is that the locational selection of human activity in the ancient times was related to the current period environmental and geographical conditions. Based on these conditions that characterize a location, repeating patterns can be identified. These, compared to patterns of other areas with similar geographic features found at the same period, may result in identifying new locations that may also have been occupied by similar human activities [1]. Namely, predictive modelling aims in establishing a causal relationship between certain environmental parameters and known archaeological site locations, in order to create a statistical model based on that relationship that can be applied to unsurveyed areas.

Taking under consideration the context of this research field of technological applications in Archaeology this study examines the geographical location of Macedonians tombs in Northern Greece (Prefectures of Macedonia and Thrace, Fig. 1) and presents a methodology for developing predictive models for detecting/identifying areas of possible Macedonian tombs’ occurrence. The study is based on documented archaeological data and geographical factors, geospatial analysis, predictive modelling and fuzzy logic. The methodology was developed upon the assumption that the locational selection of those tombs was not random, but it was a result of logical decisions based on geographical considerations and factors related to human activity (settlements, roads).
By taking into account their common spatial characteristics and documented archaeological evidence, in our study we have designed, built and implemented a predictive model to detect possible Macedonian tombs’ locations, which could benefit both archaeological research (discovery and study of new tombs, contribution in ancient topography etc.) and cultural resource management (CRM) and protection. In the following sections, we present applications of predictive modelling for burial sites, and we describe the proposed methodology and the model development process along with experimental results and evaluation of the model.

3. Archaeological predictive modelling for burial sites

Numerous archaeological predictive models have been developed to date to detect remains of human activity in the past [1–10]. However, an extensive review of the related literature indicates the lack of applications related to cemeteries and burial sites. Undoubtedly, burial mounds, tombs and cemeteries have been the subject in many studies, which, however, examine the correlation between topography and their location on the landscape [11,12], chronological estimations [13], viewshed and visibility [14–17] or simply included among other archaeological data, the locations of funerary monuments and cemeteries to map archaeological sites. The studies found in literature regarding exclusively the prediction of burial monuments or mounds are rare [18–21].

The noticeable lack of that kind of predictive models attracted our attention in the particular scientific field and triggered the effort of creating a predictive model for the detection of Macedonian tombs.

4. Model-building process

The methodology used to create the predictive model was based on the following procedures (Fig. 2): Through the archaeological research on Macedonian tombs and the data aggregation, assumptions related to their location were formulated, resulting in the selection of criteria considered to have influenced the sitting of the Macedonian tombs.

The model was created and tested under various combinations of parameters related to the criteria. The results were evaluated by using a commonly used predictive gain, which tested the predictive ability of the model for the different cases. In the following paragraphs the basic steps of the modeling process are being discussed.

4.1. Archaeological research – Data selection

The first concern for the development of the model was to collect all data regarding the burial monuments that have been discovered to date in Northern Greece, which the archaeological research entails and accepts as “Macedonian tombs”. This stage included an extensive archaeological literary research, followed by field survey, in order to locate the tombs in the areas indicated by the literature references. The total list resulted from the literature research for the area of interest (Macedonia-Thrace, Greece) included 95 Macedonian tombs. However, it was not possible to locate and acquire the geographical coordinates of all tombs, as some of these monuments are today damaged or even lost. Therefore, the tombs eventually gathered from field survey and used in this study are 83.

4.2. Selection of the criteria

The data used to create an archaeological predictive model always arise from the relationship of archaeological sites with the
natural and cultural environment. It is clear, however, that the input parameters of an archaeological predictive model should be associated both with the study area and the subject of study. In reference [22], for example, it is reported that many studies, which examine the locational processes of ancient settlements (both before and after the introduction of GIS techniques to Archaeology), suggest that, apart from socioeconomic factors, features such as topographic relief, distance from water bodies or soil cover type, had also an important role [23–32]. Those features, however, cannot be used as input data on predictive models for other types of archaeological sites (for example burial mounds or sanctuaries). Therefore, in any case, it is necessary to study thoroughly the particular type of archaeological site and extract the criteria that led to the specific human decision rules. It is clear that those factors-criteria can vary even for the same type of archaeological site, as they may be related to a specific time period, region or specific cultures.

In the case of Macedonian tombs the selection of the criteria was determined by the environmental and cultural factors that were considered to have influenced the choice of their location. The orientation, for example, of those funerary monuments could not be used as a criterion, because the archaeological research does not identify a certain pattern or certain principles [33,34]. Similarly, the distance from sanctuaries or temples could neither be used, because the archaeological research to date does not substantiate any relationship between Macedonian tombs and religious places. Thus, by taking under consideration the literature research on all Macedonian tombs, and, also, based on the existing geographic data, we ended up with four environmental (altitude, slope, soil hardness, distance from rivers) and two cultural parameters (distance from settlements, distance from roads).

4.3 Predictive model

The stages of a multicriteria analysis, also followed in the present predictive model, include: formation of the problem, quantification of the criteria, calculation of the criteria weights and aggregation of the criteria (Fig. 3). These four stages are discussed in the following paragraphs.

4.3.1 Formation of the problem

The spatial distribution of known Macedonian tombs in Macedonia and Thrace covers a large area (about 42,770 km²), with many of them being concentrated around ancient settlements (Pella, Vergina etc.), while others being isolated in various locations from Kozani (West Macedonia) to Orestiada (East Thrace). The survey of such a large area, as expected, impedes the implementation of a model for the prediction of exact Macedonian tombs’ locations. On the other hand, focusing only on a particular area within Macedonia and Thrace, would be incorrect, as Macedonian tombs have been discovered scattered across these two geographic regions of northern Greece. Therefore, the study had to be oriented towards the development of a predictive model that would specify areas (and not exact locations) of high probability of Macedonian tombs’ occurrence.

4.3.2 Quantification of the criteria

A prerequisite for the implementation of the proposed predictive model was to quantify the selected criteria. The first stage of this process included the specification of certain areas (buffer zones) around dimensionless data, such as vector map data (points and lines). In order to specify the range of the buffer zones, namely the size of the space formed around the linear and point data, a statistical analysis of the distances of the known Macedonian tombs from settlements, roads and rivers was conducted. The process resulted in the use of the median value and its multiples.

In the second stage of the quantification process, all criteria were normalized so that they could be referenced in a common scale, and therefore the criteria aggregation would be performed on a common basis. To this end, we normalised the criteria values using fuzzy logic using linearly ascending/descending membership functions, where each criterion was normalized into a byte range (0 to 255), where 0 represents the minimum and 225 the maximum probability of finding Macedonian tombs.

4.3.3 Selection of importance and calculation of criteria weights

The third stage of the model-building process included the calculation of the criteria weights, namely the importance of each criterion. One of the most popular methods used in a multicriteria analysis is the Analytical Hierarchy Process (AHP) [35], a structured technique for organising and analysing complex decisions. In AHP one can attribute different weights to the criteria or sets of activities, depending on the degree of their significance, by making pairwise comparisons of the criteria, based on the decision makers’ judgments about their relative meaning and importance.

4.3.4 Criteria data aggregation

The process of criteria data aggregation was achieved by using the method of weighted linear combination [36], whereby each
criterion’s value is multiplied with the value of its weight and the results are summed. Provided that the sum of all weights equals to 1, the result of the aggregation will have the same range as the one specified for the criteria, which in this case, as mentioned above, ranges from 0 to 255. The aggregation process can be mathematically expressed by the following equation:

$$ S = \sum w_i x_i $$  

(1)

where $S$ is the probability of Macedonian tomb occurrence, $w_i$ the weight of the criterion $i$ and $x_i$ the value of the criterion $i$.

5. Experimental results and model validation

The proposed predictive model has initially been tested for cases where all six criteria were considered equally important and therefore shared the same weight. Additionally, by taking into account the archaeological research that documents the existence of Macedonian tombs near settlements and roads, the criteria “distance from settlements” and “distance from roads” were considered to be of higher importance than other criteria, and so they could be assigned higher weights. This could be quantified by varying the significance of these criteria using a multiplier ($2x$ to $6x$).

Thus, in order to examine the predictive capability of the model, the uncertainty regarding the selection of the range of buffer zones and the sensitivity of the model, we have tested the model using a total of twenty-seven ($27$) combinations of the criteria weights ($w$) and the buffer zones’ range ($m$) that resulted from multiples of the various distances median value. The $27$ different scenarios of the model resulted in the production of $27$ probability maps. Table 1 shows the selected combinations.

In Figs. 4–6, we present results of three tests of the model. In the generated maps the range 0–255, is expressed in five (5) colored shadings, each one of which corresponds to an area of specific probability. On these maps we overlap the known Macedonian tombs’ locations as black dots, in order to provide a first impression about the predictive efficiency of the model.

The simple, visual evaluation of the results, indicates that a large number of Macedonian tombs are located in areas of high and very high probability for the scenario $w1/m1$ (Fig. 4), where the buffer zones equal to the median of the distances of the known Macedonian tombs from settlements, rivers and roads. The results of the scenarios $w3/m1$ and $w5/m1$ (Figs. 5 and 6) suggest that higher weight values of the criteria “distance from settlements” and “distance from roads” significantly reduce the area of high and very high probability zones and thus affect the identification of the known Macedonian tombs in all areas.

A first conclusion that emerges from the initial processing of all the test results is that an increase in the buffer zones ($m$) leads to an increase in the number of known tombs identified in the high and very high probability areas. The second conclusion is that the differentiated weight values of the “distance from settlements” and “distance roads” criteria influence the predictive ability of the model up to a certain extent. Therefore, there is no point in increasing the weights for those two criteria, above a certain value, as it is shown, in many tests, that the increased weight values affect very little both the spatial identification of Macedonian tombs and the area of the high and very high probability zones (Figs. 7 and 8). Again, an important result regarding the predictive ability of the model is that in all of the $w1$ and $w2$ tests none of the known tombs is identified in a very low probability area.

However, the fact that it has been possible to construct a predictive model does not in itself guarantee the accuracy of its predictions [37]. The validation of the model must be examined with reference to the areas, which the model indicates as most likely to find Macedonian tombs. For example, a model can successfully predict 80% of the sites, within a region of 70% of the total area. This model practically rules out only a small part of the survey area, so it cannot be considered as an efficient one. Kwanne [38] proposed the validation of a predictive model, defining a predictive gain as follows:

$$ G = 1 - E_1/E_2 $$

where, in the case of our model, $E_1$ is the percentage of the total area where tombs are predicted, and $E_2$ is the percentage of known tombs identified within the area where they are predicted, for a given probability of site occurrence.

It is noted that the gain $G$ is calculated for a specified probability of archaeological sites occurrence. The gain ranges between $1$ and $1$, where a zero (0) value indicates no predictability, value one (1) indicates the highest predictive utility and minus one ($-1$) the highest predictive utility for the reverse of what it is supposed to.

In order to validate our model using the gain $G$, we used the spatial coordinates of all known Macedonian tombs that have been collected. In an attempt to identify the areas of high and very high probability with a high allocation of known Macedonian tombs we have isolated the results of those tests that detect more than 75% of the known tombs (Table 2). The maximum value of gain $G$ (0.812) is achieved for $w2/m1.5$, where 87.95% of the known Macedonian tombs fall in an area of 16.55% of the total survey area. On the contrary, the minimum value of gain $G$ (0.435) resulted for $w1/m3$, in which a great percentage (97.59%) of the tombs was found, but in a large land area (55.16%). The latter result indicates that there is a threshold in increasing the buffer zones. Any higher value above

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Combinations of the criteria weights ($w$) and the buffer zones’ range ($m$).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria weights ($w$)</td>
<td>Median multipliers ($m$)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$w1/m0.15$</td>
</tr>
<tr>
<td>2</td>
<td>$w2/m0.5$</td>
</tr>
<tr>
<td>3</td>
<td>$w3/m0.5$</td>
</tr>
<tr>
<td>4</td>
<td>$w4/m0.5$</td>
</tr>
<tr>
<td>5</td>
<td>$w5/m0.5$</td>
</tr>
<tr>
<td>6</td>
<td>$w6/m0.5$</td>
</tr>
</tbody>
</table>

Please cite this article in press as: A. Balla, et al., Locating Macedonian tombs using predictive modelling, Journal of Cultural Heritage (2012), http://dx.doi.org/10.1016/j.culher.2012.10.011
that threshold does not improve the predictive ability of the model (even though a large percentage of Macedonian tombs is detected), as it also increases the area of high and very high probability zones. Therefore, the best performance of the predictive model lies in the use of $w2/m1.5$, obtaining a high percentage of Macedonian tombs in just 16.55% of the total land area of interest.

Fig. 9 shows graphically the distribution of known Macedonian tombs in the areas of high and very high probability, the areas of the corresponding zones and the gain $G$ for those tests that the correct detection of the tombs is more than 75%.

An interesting challenge was to study the behaviour of the model in cases where the original problem is inversed, and the aim now is to identify the largest possible areas that do not include tombs. In such cases, predictive models can be used to exclude areas that are expected to contain Macedonian tombs from any development/urban planning, in order to prevent possible future damage of those monuments by modern development interventions. The answer to the inverse question should be inquired in the results of the tests referring to low and very low probability zones and moreover in those tests with total tombs absence.

Fig. 4. Map results showing probability zones and distribution of known Macedonian tombs for the case study $w1/m1$.

Fig. 5. Map results showing probability zones and distribution of known Macedonian tombs for the case study $w3/m1$.

Fig. 6. Map results showing probability zones and distribution of known Macedonian tombs for the case study $w5/m1$. 

Please cite this article in press as: A. Balla, et al., Locating Macedonian tombs using predictive modelling, Journal of Cultural Heritage (2012), http://dx.doi.org/10.1016/j.culher.2012.10.011
Fig. 7. High and very high probability zone area as a function of the criteria weight for different buffer zone ranges.

Fig. 8. Detected known tombs in high and very high probability zones as a function of the criteria weight for different buffer zone ranges.

Fig. 9. Detected known tombs in high and very high probability zones, the area of these zones and the gain $G$ for the cases where the detection of known tombs is more than 75%.
The validation of the predictive model for the inverse problem is analogous to the evaluation process of the cases presented above (high and very high probability). However, it should be noted, that in cases of low and very low probability the gain G cannot be defined, as the number of tombs found in these areas is zero. Therefore, to evaluate the results of the model it would be sufficient just to acknowledge the largest possible area, in which none of the known tombs is found. Table 3 shows the results of the tests, where no known tomb was found in low and very low probability areas. The best performance of the model regarding the inverse problem is attained for w1/m1, where the highest percentage of land area of the total survey area was achieved (29.60%). The least satisfactory results are those attained for w1/m3, where only 10.94% of the total area can be excluded.

Fig. 10 represents the percentage of area of low and very low probability zones, when no known tomb is found.

6. Conclusions

This paper described a general methodology used to create a predictive model that can be applied to identify probability zones of Macedonian tombs’ occurrence. The model built on a hybrid approach, taking into account cultural and geospatial factors based on extensive archaeological research. All known and identifiable Macedonian tombs were studied and their actual locations were recorded in order to serve as ground truth data. The evaluation of the model demonstrated the validity of the proposed methodology. The prediction results, in the form of quantized prediction maps, are considered satisfactory as they provide answers to a series of questions that are related to the problem at hand (archaeological research, cultural resource management and protection, land use, etc.).

In the case where the main objective was the contribution to archaeological research, the goal was to identify a large percentage of known Macedonian tombs into the most confined areas possible. In this sense, the model can contribute to a possible cost reduction by minimizing the requirements for trial excavations, as it indicates areas that are of high probability to crown an archaeological excavation with success. The model provided very promising results, identifying a high percentage of known Macedonian tombs within relatively small spatial zones.

On the other hand, the results produced by the model when used for cultural heritage protection (exclusion of large areas with no probability of archaeological sites’ occurrence), were also important. The model in this case can be considered as a useful tool that can be applied in decision making about use of land and modern development activities, as it can tackle with such problems by indicating relatively large spatial zones with absolute absence of archaeological locations.

In addition, the methodology adopted in this work could be generalized as a workflow applicable to archaeological predictive modelling for other types of archaeological sites, as long as there is a strong geospatial “character” in the involved data. By conducting archaeological research on a specific type of archaeological site and by customising the input data of the model (criteria, geospatial features related to the site), the adopted methodology could also result in successful predictive modelling for other types of archaeological remains.

References


Please cite this article in press as: A. Balla et al., Locating Macedonian tombs using predictive modelling, Journal of Cultural Heritage (2012). http://dx.doi.org/10.1016/j.jculher.2012.10.011