RECENT ADVANCES IN ARCHAEOLOGICAL PREDICTIVE MODELING FOR ARCHEOLOGICAL RESEARCH AND CULTURAL HERITAGE MANAGEMENT

Aikaterini Balla¹, Gerasimos Pavlogeorgatos², Despoina Tsiafakis³
George Pavlidis³

¹ 31st Ephorate of Prehistoric and Classical Antiquities, Archaeological Museum of Abdera
67061, Abdera, Xanthi, Greece
² Department of Cultural Technology and Communication, University of the Aegean, Mytilini,
Greece
³ Institute for Language and Speech Processing, ATHENA Research Centre, Greece

Received: 01/12/2013
Accepted: 17/07/2014

Corresponding author: G.Pavlidis (gpavlid@ceti.gr)

ABSTRACT

The identification of areas that are insignificant for archaeological research can be used for guidance and support in projects that involve decision-making about the use of land and modern development activities. On the other hand, the identification of areas significant for archaeological research can contribute to archaeological knowledge and minimise the risk of unsuccessful excavations.

This paper presents a review of the most recent and representative applications of predictive modelling in Archaeology, which demonstrate that predictive models can be successfully exploited by archaeological research and Cultural Heritage Management (CHM).

KEYWORDS: archaeological predictive modelling, Cultural Heritage management, archaeological research
1. INTRODUCTION

The primal objective of Archaeology is the composition of the history and the understanding of past cultures through the study and interpretation of the natural relations of archaeological finds and the ideological context within which they operate. The problem, however, is always more complex in practice. On one hand, the discovery of archaeological remains is coincidental and mainly a result of modern development interventions, which, however, lead to partial or total destruction of archaeological sites. On the other hand, a number of reasons, such as the lack of financial resources ensure lack of systematic archaeological excavations and therefore, incomplete knowledge of the archaeological remains in many areas. Even in the cases of more systematic excavations, the studied areas are usually necessarily small and, consequently, the archaeological information collected and studied cannot be easily compared or opposed to data from other areas related to the same human activity, from which these archaeological remains were generated.

The effort to address these problems led to the development and implementation of methodologies that would be able to recognize and identify possible areas of human activity and use in the past. Within this research context, the use of archaeological predictive modelling in the past years has yielded important expertise that can be used successfully both in CHM and archaeological research.

Management of cultural heritage over the recent decades received more attention and resources in order to face threats related to the physical damage of cultural assets during modern development activities (urban development, large-scale agriculture and mining), or even threats like looting and environmental threats like erosion (Neumann and Sanford 2001). Recently, the identification and protection of cultural sites has been a subject of legislation that in many cases criminalised land development prior to conducting a cultural resources survey to identify any cultural sites that may be affected (Neumann and Sanford 2001). Nowadays, CHM can be significantly empowered by the application of scientific approaches and digital technologies that provide fast and reasonably accurate prediction of the existence of cultural sites prior to development projects. Predictive Modelling (PM), in particular, has already been proven as a valuable tool for the rescue of archaeological remains and archaeological data that would otherwise have been lost due to modern development.

Additionally, in terms of archaeological research, PM has already been successfully exploited by archaeologists, and is further expected to be an integral part of archaeological practice, in interpreting and understanding the socio-economic structure of the past. The identification of new archaeological sites through the application of PM techniques would enrich archaeological knowledge about ancient culture and would contribute to the study of ancient topography, as the discovery of new sites can result in finding yet undiscovered areas of archaeological interest. Furthermore, the predictive models can be used as an efficient solution to the lack of funding, by providing insight on the existence of archaeological remains in studied areas, thus minimising the need for trial excavations.

2. ARCHAEOLOGICAL PREDICTIVE MODELLING

The most commonly used definition of PM in Archaeology belongs to Kohler & Parker (1986), who describe it “a technique that, at a minimum, tries to predict the locations of archaeological sites or materials in a region, based either on a sample of that region or on fundamental notions concerning human behaviour”.

According to Verhagen (2007), PM is based on the assumption that the location of archaeological sites is not random, but it is associated to specific characteristics of the natural environment and factors related to human activity and human behavioural norms in the past. By identifying this causal relationship between certain environ-
mental and geographical characteristics and known archaeological site locations, repeating patterns can be identified, creating a statistical model that can be applied to unsurveyed areas in order to identify new locations that may also have been occupied by similar human activities. Namely, PM can be conceptualised as a specialised form of location-allocation analysis, where the aim is to allocate suitable locations to specific types of human activity and their archaeological remains (Van Leusen 2002).

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. Jaroslaw and Hildebrandt-Radke (2009) for example, report 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 (Bauer et al. 2004, Duke 2003, Fletcher 2008, Kvamme 1992, Stancic and Kvamme 1999, Warren 1990, Willey 1953, Williams 1956, Williams et al. 1973). 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.

3. CONTRIBUTION TO CHM AND ARCHAEOLOGICAL RESEARCH: INDICATIVE CASE STUDIES

The increased attention in archaeological PM led to an extensive literature research and numerous case studies. In this study, we present indicative studies of archaeological predictive modelling of the past ten years, analyse their aims and scopes to both CHM and archaeological-academic research and present their experimental results.

Siart et al. (2008) conducted geospatial analyses of archaeological sites and communication paths of the Bronze Age (Minoan Neo-Palace Period - about 1650 BC) in the region of Mount Ida in central Crete, for the detection of Bronze Age infrastructures and potential archaeological candidate sites. The study included the development of an information system which visualized the main geological characteristics of Mount Ida, the mapping (based on field research) of the geomorphology, the vegetation, the hydrology and known archaeological sites of the study area, remote sensing techniques, least cost analyses, predictive modelling and GIS. The researchers stress the need to include in archaeological analyses comprehensive sets of environmental variables that might have influenced ancient settlement patterns and show the advantages of using a multi-method approach for reconstructing ancient landscapes. The study can be used, according to the researchers, for unexplored areas where the archaeological data are poor and the environmental conditions adverse and can contribute to the acquisition of new archaeological knowledge.

Vaughn and Crawford (2009) used GIS in conjunction with remotely sensed imagery, paper map data, and Binary Logistic Regression to predict the probability of ancient Maya archaeological site presence in Northwest Belize in Central America. The input variables in the modelling process were selected using Binary Logistic Regression and were associated with both the representation of the ancient landscape and the current environmental terrain. The optimal function of the model was achieved by the use of the following variables: vegetation index (Tasselled Cap Greenness index), viewshed (eastern aspect), proximity to flat land. The model suggested that there is a higher probability of settlements’ oc-
currence in locations oriented to the east, with easy access to arable land and high vegetation index. The evaluation of the model was examined using the Kvamme predictive gain $G$, which generated a moderate gain statistic of $0.26$. The proposed methodology can be used, according to the researchers, to the archaeological-academic research regarding the Mayan culture, and moreover, greatly reduce the cost and time required for future field research.

Fernandes et al. (2011) used statistical methods and GIS to study the settlement site in Malia in the ProtoPalatial period and understand the causal relationship between settlement locations and independent variables. They applied two predictive models: a purely environmental, which used as criteria-variables altitude, distance from sea coast, geomorphology, soil depth and density of the water bodies and a mixed model (environmental-historical), which used, apart from most of the environmental variables mentioned above, criteria concerning human factors such as the major urban centres of the time. The researchers introduced in the modelling process an algorithm in order to identify the best performance of the model. The testing and evaluation of the results of the two models ascribed the best predictive ability in the environmental-historical model.

Graves (2011) developed and applied two predictive models in order to identify human settlement and occupation activity in the mainland of Scotland during the Neolithic Period. The study was based on GIS, statistical methods and an inbuilt presumption that locations of settlement or occupation activity on the mainland were related to the locations of the timber halls, pits, and chambered cairns. A GIS was used to extract environmental variables commonly used in archaeological predictive modelling from input sites and non-site locations: the variables included elevation, slope, aspect, local relief, distance to the nearest source of water, cost-distance to the nearest source of water, and viewshed. An important conclusion of the study was that the variables of viewshed and proximity to water bodies seem to have greater significance in the selection of the sites in relation to the other criteria as they were identified as powerful predictors. To evaluate the prediction results Graves used gain $G$ which showed that, out of a total of 74 ‘activity’ sites, models 1 and 2 can successfully predict 86% and 84.5%, respectively. However, as Graves points out, without fieldwork it is impossible to know the real gain of each model and therefore the gains should be treated as preliminary and secondary to field tests. The researcher finally fosters the hope that in the future, the proposed models could test in the field the archaeological theories about the perceived relationship of the input sites to settlement or occupation activities.

The study subject of Aubry et al. (2012) was the open-air rock art of the late Ice Age and the Iron Age, which shows a similar spatial distribution along the rivers Côa and Douro (Portugal) and orientation towards Southeast. The researchers tried to determine whether the artists of the two periods deliberately chose the same natural environments for their art, or if the current spatial distribution of their art remains resulted from other processes formations or corrosion of rocks, before or after the artistic formations. The study included analysis of geological structures of local and regional level, field measurements, analysis of the hydrological network, etc., whereby the researchers concluded that the distribution of the remains of prehistoric art is the result of natural processes and geological formations combined with different conservation/erosion of the rocks’ surface. The factors that affected, according to the researchers, the erosion of these surfaces were the diversity of solar radiation, humidity, and the growth of algae and lichen. The researchers combined the interpretation of observations from field research with probabilistic processes, pairwise comparisons of observed patterns of the two time periods and geospatial analysis through GIS to develop predictive models, which would identify areas with similar geological and climatic characteristics. The
archaeological input data (rock art occurrences) were used to evaluate the predictive models and external validation maps, with the results showing an agreement of 70%-80%. Most importantly, the following field survey, revealed unknown rock paintings in areas with high and very high values. The researchers conclude that the predictive model would serve as a useful tool in archaeological fieldwork and Cultural Resource Management.

Luczak (2013) used two regression modelling methods: Generalized Linear Model (GLM) and Generalized Additive Model (GAM) and examined their ability to analyse and predict archaeological sites locations in southern Poland. The archaeological datasets used in the modelling process came from the field survey record stored in the database of the Polish Archaeological Record (PAR) and represented sites from two different periods: Neolithic and Medieval (Early and Late). Through typical GIS software and procedures 11 environmental variables (representing hydro-morphological terrain attributes and soil types) and a set of cultural variables (visibility and distance from political and administrative centres (castle, fortress or fortified settlement) were obtained and statistical analysis (e.g. density plots, correlation coefficients, etc.) was used in order to determine past settlement preferences, their potential influence on site location and also to examine the differences between settlement patterns in these periods. The models’ predictive ability was evaluated through the ROC curve function (AUC), which showed that statistically the GAM models give better predictions than the GLM models. Łuczak chose the best modelling method (the GAM models) to produce binary probability (0-1) maps, which were next used to create 2 final maps for Neolithic and Medieval periods combining permanently and temporarily settled sites predictions. The researcher concluded that predictive models could serve as a great tool for archaeologists in settlement research, but also stresses the need to be verified and checked for their reliability, apart from their statistical evaluation, through field surveys. Moreover, it is necessary to use a variety of prediction methods to understand and interpret different results and also examine the accuracy of the models not only for chronologically different sites (prehistoric and historical), but also for different site types (e.g. temporarily settled, permanently settled, hunter shelter, monuments etc.).

Verhagen et al. (2013), noting the relative absence of socio-cultural factors in prehistoric and historical site location choice attempted to address the unexplored human factor in predictive modelling, by developing a protocol using both environmental and socio-cultural factors that can easily be implemented for different regions and time periods. The development of the predictive model was based on cross-regional comparisons of settlement location factors, like slope, aspect and solar radiation made in the 1990s by analysing the environmental context of Roman settlements in the French Rhône Valley. However, for the current study, the researchers expanded the set of variables with “socio-cultural” factors, in particular accessibility, visibility, and the effect of previous occupation in order to establish whether including socio-cultural factors actually made a difference for the interpretation of site location patterns and predictive model quality. Though the prediction of settlement locations was implied, the researchers stressed that the optimal model performance was not the main goal of their study, as would be the goal of standard statistical approaches like logistic regression. Instead, the “non-performance” of a variable was considered an equally important result, as the protocol’s aim was to extract the main factors that influenced settlement location over a longer term. In conclusion, the study, though preliminary, showed that there were limitations to both environmental and certain social variables, as they may not be relevant for other archaeological settings, or cannot even be modelled in all situations because of poor available archaeological data.

An extensive review of the related literature indicates the lack of applications relat-
ed to cemeteries and burial sites. Undoubt-
edly, 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 (De Reu et al. 2011, Löwen-
borg 2010b), chronological estimations
(Löwenborg 2009), viewshed and visibility
(Fisher et al. 1997, Lageras 2002, Wheatley
1995, Woodman 2000) or simply included
among other archaeological data, the loca-
tions of funerary monuments and cemeter-
ies to map archaeological sites. The studies
found in literature regarding exclusively
the prediction of burial monuments or
mounds are rare and will be presented
shortly in the following paragraphs.

Al-Muheisen & Al-Shorman (2004) used
GIS to analyse the landscape and the mor-
tuary practices in three cemeteries of the
late Roman and early Byzantine period in
the region Bediyeh (North Jordan), in order
to obtain the spatial relationships of the
various features at the site in a ritual and
cultural context and, thus, reconstruct past
behavioural practice and cult. Within the
framework of their research they devel-
oped an inductive predictive model for
burial monuments and applied it to the ar-
chaeological data of the western cemetery,
which, based on the typology of the tombs
(chambered) and the funerary gifts found
inside the tombs, was believed to be pre-
destined for the dead of the higher social
classes. In the model building process they
used three variables (proximity between
the funerary monuments, slope, viewshed),
to which a different weight was assigned,
based on the frequency of the known mon-
uments in relation to each of these criteria.
The prediction results were evaluated with
the gain G, attaining 0.82 value for the best
performance of the model. It was noted
that the locations indicated by the model as
the most probable for tomb occurrence
were places prominent and visible to other
tombs, which can be attributed, as the re-
searchers speculate, to the higher social
class of the deceased and their relatives’
desire for their tombs to be visible. Apart
from the contribution in the understanding
of the various spatial relationships among
the various features of the site and inter-
pretation of the archaeological data, the
predictive model can also be used as a
guide and, moreover, as a “cost reducing”
tool for future excavations.

Fry et al. (2004) within the context of
Cultural Heritage Management and protec-
tion in Norway suggested a methodology
for the development of predictive models
that would be able to identify possible loca-
tions of burial monuments’ of the Bronze
and the Iron Age. For the spatial analysis
and the mapping of the spatial data they
used GIS, readily available environmental
data and visual analysis. The predictive
model indicated areas of high probability
of burial mounds’ existence and success-
fully provided 94% of the known mounds in
areas that cover only 12% of the total sur-
vey area. The results led to the discovery of
new sites of archaeological interest and
have contributed significantly to the un-
derstanding of the tomb distribution in
Norway. Moreover, the generated maps
would serve as a useful tool for heritage
managers and development projects’ plan-
ners by identifying areas where develop-
ment may risk damaging antiquities.

Burns et al. (2008) have created models
that predict possible locations of funerary
monuments in the Theban necropolis in
Luxor, Egypt, and examined its usefulness
in understanding the reasons that led to the
preference of those locations by the ancient
Egyptians. The location of the tombs was
examined in relation to geology, slope, ele-
vation, fractures, and religious-funerary
practices (orientation of tombs, proximity
to temples). The environmental and the ar-
chaeological data were quantified using
GIS and statistical analyses and a predi-
cptive model was developed, which could be
linked to the database of the Egyptian Ant-
quities Information System (EAIS) created
by the Egyptian government for the protec-
tion of Cultural heritage in order to indi-
cate which sites should be avoided within
the context of modern development plan-
ning or further studied in terms of archaeo-
logical research. The model results showed
that the stable soil, the eastward orientation and the orientation to the royal temples were important factors in the decision-making process regarding the siting of the tombs. These models, however, did not allow determining the degree of importance of these factors. On the other hand, it seems that the criteria of elevation, slope and fractures did not affect significantly the choice of the tombs’ location. However, according to the researchers, this does not mean that the ancient Egyptians did not take them under consideration, but that their study obviously was lacking relevant archaeological documentation. The researchers suggested for future study the examination of the role of other factors that may have affected the tombs’ location. The most important outcome of their study is related to the lack of archaeological knowledge about the reasons-factors that led to the locational selection of the tomb construction. Most archaeological studies are limited to the discovery of the identity of the deceased and not necessarily to the knowledge of the criteria, which led to the construction of the burial monument at this particular site. Further investigation of this field of Egyptian Archaeology could lead to better models and thus to the discovery of new tombs and the understanding of the complexity of decision making in the past.

Balla et al. (2013) created a predictive model for the detection of Macedonian tombs in Northern Greece. The proposed methodology was based on the following procedures: through archaeological research and data aggregation, assumptions related to the location of the sites of interest were formulated, resulting in the selection of criteria considered to have influenced the siting of the Macedonian tombs. Thus, by taking under consideration the literature research on all Macedonian tombs, and, also, based on the existing geographic data, the researchers ended up with four environmental (altitude, slope, soil hardness, distance from rivers) and two cultural parameters (distance from settlements, distance from roads). At the core of the proposed methodology, a multi-criteria analysis on geospatial data processing technologies (GIS), predictive modelling techniques and fuzzy logic was 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 created and tested under various combinations of parameters related to the criteria. The results were evaluated by using a commonly used predictive gain, which proved the efficiency of the model’s predictive ability in providing answers to a series of questions related to the problem at hand and could benefit both archaeological research (discovery and study of new tombs, contribution in ancient topography etc.) and cultural heritage management and protection. On one hand, in terms of archaeological research, the model provided very promising results, identifying a high percentage of known Macedonian tombs within relatively small spatial zones. Namely the model identified a total of 87.95% of the known Macedonian tombs within a 16.55% of the total survey area and in the case of the “very high probability” areas, identified 55.42% of the known Macedonian tombs in an area smaller than 6% of the total surface area (namely in the 1/19 of the total survey area). On the other hand, the results produced by the proposed method were considered of great importance for cultural heritage protection. In that case, where the aim was the identification and knowledge of large areas that contain no, or the least possible, archaeological sites, the model indicated a large area with total Macedonian tombs’ absence (31.73% of the total survey area), which could be excluded in development planning.

Table I summarises the recent works in archaeological predictive modeling presented in this paper.
Table I. Summary of recent works in archaeological predictive modeling

<table>
<thead>
<tr>
<th>Authors</th>
<th>Prediction Aim</th>
<th>Area/period of interest</th>
<th>Tools</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siart et al. (2008)</td>
<td>Bronze Age infrastructures</td>
<td>Region of Mount Ida in central Crete (Bronze Age)</td>
<td>Remote sensing, GIS</td>
<td>Contribution to archaeological research and CHM</td>
</tr>
<tr>
<td>Vaughn and Crawford (2009)</td>
<td>Ancient Maya archaeological site presence</td>
<td>Northwest Belize in Central America</td>
<td>GIS, remote sensing, paper map data, binary logistic regression</td>
<td>Kwamme predictive gain 0.26, discovery of new sites, contribution to archaeological research</td>
</tr>
<tr>
<td>Fernandes et al. (2011)</td>
<td>Causal relationship between settlement locations and environmental and cultural variables</td>
<td>Malia, Crete (Proto-Palatial period)</td>
<td>Statistical methods and GIS</td>
<td>Better performance by using environmental-historical model</td>
</tr>
<tr>
<td>Graves (2011)</td>
<td>Human settlement and occupation activity (Neolithic Period)</td>
<td>Mainland of Scotland</td>
<td>GIS, statistical methods and specific limitations</td>
<td>Successfully predicted 86% of 74 sites, contribution to archaeological research</td>
</tr>
<tr>
<td>Aubry et al. (2012)</td>
<td>Factors that influenced the location of open-air rock art in different time periods</td>
<td>Rivers Côa and Douro, Portugal (late Ice Age and the Iron Age)</td>
<td>Probabilistic processes, pairwise comparisons and geospatial analysis through GIS</td>
<td>70%-80% prediction accuracy and new sites identification, contribution to archaeological research</td>
</tr>
<tr>
<td>Luczak (2013)</td>
<td>Identification of Neolithic and Medieval (Early and Late) settlements</td>
<td>Southern Poland (Neolithic and Medieval - Early and Late)</td>
<td>Generalized Linear Model (GLM) and Generalized Additive Model (GAM)</td>
<td>New sites identification, contribution to archaeological research</td>
</tr>
<tr>
<td>Verhagen et al. (2013)</td>
<td>Protocol using environmental and socio-cultural factors that can easily be implemented for different regions and time periods</td>
<td>Roman settlements in the French Rhône Valley</td>
<td>---</td>
<td>Extracted the main factors that influenced settlement location</td>
</tr>
<tr>
<td>Al-Muheisen &amp; Al-Shorman (2004)</td>
<td>Burial sites identification</td>
<td>Region Bediyeh (North Jordan)</td>
<td>GIS</td>
<td>Kwamme predictive gain 0.82, discovery of new sites, cost reduction, contribution to archaeological research</td>
</tr>
<tr>
<td>Fry et al. (2004)</td>
<td>Burial sites identification</td>
<td>Norway (Bronze and the Iron Age)</td>
<td>GIS, readily available environmental data and visual analysis</td>
<td>94% accuracy in 14% of total survey area, discovery of new sites, contribution to archaeological research</td>
</tr>
<tr>
<td>Burns et al. (2008)</td>
<td>Burial sites identification</td>
<td>Theban necropolis in Luxor, Egypt</td>
<td>GIS and statistical analysis</td>
<td>Contribution to archaeological research and CHM</td>
</tr>
<tr>
<td>Balla et al. (2013)</td>
<td>Burial sites identification (Macedonian Tombs)</td>
<td>North Greece (Late Classical and Hellenistic)</td>
<td>Multi-criteria analysis GIS and fuzzy logic</td>
<td>Probability maps, contribution to archaeological research and CHM</td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

Predictive models for archaeological sites have become an integral part of archaeological applications, displaying an increasing number of methodologies that attempt to meet different purposes and needs of Archaeology (contribution to archaeological research or CHM). Despite the differences in the analysis approach, practically, the same process is followed in all cases: their creation is based on the correlation of environmental and cultural parameters with known archaeological sites. The statistical analysis of those archaeological sites correlates the spatial variability of the environmental and cultural parameters with other sites of possible archaeological interest, based on specific decision making rules.

Predictive models can be used in archaeological-academic research by indicating
areas of high probability to find sites of archaeological interest and therefore need further investigation. Thus, the discovery of new archaeological sites would certainly add new data to the existing archaeological knowledge and the study of the historical topography, providing a clearer picture of the number of sites of human activity in the past, their spatial relationships, their connecting networks (roads) etc. Additionally, they can contribute to a cost reduction by minimising the requirements for trial excavations.

On the other hand predictive models can be used for cultural heritage protection, where the aim is to identify the areas that do not include sites of archaeological interest and, thus, exclude them from any development/urban planning. When used as a spatial guidance and support for projects of land use and modern development, predictive models can prevent possible future damage of archaeological sites.

In conclusion, predictive modelling can be a successful tool in archaeological analyses and studies with the potential to give new impetus to archaeological thinking and interpretation of the remains of the past.

REFERENCES


