First thoughts on the incorporation of cultural variables into predictive modelling

Verhagen, P.; Kamermans, H.; Leusen, M. van; Deeben, J.; Hallewas, D.P.; Zoetbrood, P.; Niccolucci F., Hermon S.

Beyond the Artifact

Digital Interpretation of the Past

Proceedings of CAA2004

Prato 13–17 April 2004

Edited by

Franco Niccolucci and Sorin Hermon

Franco Niccolucci and Sorin Hermon
Editors

Elizabeth Jerem
Managing Editor

András Kardos
Typesetting and Layout

Stephanie Williams
English Revision

Archaeolingua
Cover Design

Cover image: *After the Etruscan Bucchero Incenser of the Artimino Archaeological Museum*

ISBN 978-963-9911-10-9

Published by ARCHAEOLINGUA
Printed in Hungary by Prime Rate

Budapest 2010

Content

ARCHAEOLOGICAL LANDSCAPES AND GIS APPLICATIONS

QUANTITATIVE METHODS

First Thoughts on the Incorporation of Cultural Variables into Predictive Modelling

Philip Verhagen[1], Hans Kamermans[2], Martijn van Leusen[3], Jos Deeben[4], Daan Hallewas[4] and Paul Zoetbrood[4]

[1]RAAP Archeologisch Adviesbureau, Amsterdam, The Netherlands
ph.verhagen@raap.nl
[2]Faculty of Archaeology, Leiden University, Leiden, The Netherlands
h.kamermans@arch.leidenuniv.nl
[3]Institute of Archaeology University of Groningen, Groningen, The Netherlands
p.m.van.leusen@let.rug.nl
[4]State Service for Archaeological Heritage Management, Amersfoort, The Netherlands
j.deeben@archis.nl
d.hallewas@archis.nl
p.zoetbrood@archis.nl

Abstract. Predictive maps are increasingly used at all administrative levels for purposes of planning and determining policy priorities. However, current methods yield predictions with limited specificity. It is believed that methodological improvements, such as the use of non-environmental variables, will lead to a better performance of the models. The paper aims to show in what way cultural variables can actually be included in predictive modelling.
Keywords: Predictive modelling; Archaeological heritage management; Cultural variables; GIS

1. Introduction

Predictive modelling is a technique used to predict archaeological site locations on the basis of observed patterns and/or assumptions about human behaviour (Kohler and Parker 1986; Kvamme 1988, 1990). It was initially developed in the USA in the late 1970s and early 1980s where it evolved from governmental land management projects and is still regularly applied in cultural resources management. In the Netherlands, predictive modelling plays an important role in the decision making process for planning schemes on a municipal, provincial and national level.

However, in many other countries predictive modelling is far from being an accepted tool for archaeological heritage management (AHM), and even where it is used regularly, criticism is not uncommon (see e.g. Ebert, 2000; Whitley, in press; van Leusen et al., 2002). Much of this criticism is related to the uncritical application of so-called ‘inductive’ modelling techniques, in which the archaeological data set is used to obtain statistical correlations between the location of archaeological sites and environmental variables such as soil type, slope or distance to water. The performance of these models is in general not very good, partly because of the use of inappropriate statistical techniques, but mainly because of the biased nature of many archaeological data sets and the emphasis on environmental factors, which are easier to model than the more intangible social and cultural factors.

Wheatley (2003) even states that, as predictive modelling doesn’t work very well, it shouldn’t be used at all: “Archaeology should really face up to the possibility that useful, correlative predictive modelling will never work because archaeological landscapes are too complex or, to put it another way, too interesting.” His argument is mainly directed against the use of biased archaeological data sets, that will lead to the development of biased models that will in turn inevitably produce a positive feedback loop of even more biased data sets, as it is common practice to spend funds for AHM on the areas of ‘high archaeological value’. These areas will become better and better known, whereas the areas that are designated a ‘low value’ on the predictive map will largely be ignored in (commercial) archaeological research.

Verhagen (in press) however shows that the creation of biased data sets is not just a problem of predictive modelling, but a more general characteristic of the way in which archaeological data is collected. Most of the archaeological prospection done is not taking into account statistical sampling theory, and it can be suspected that many survey projects do not even have a strong archaeological hypothesis in mind. We believe that predictive modelling can serve as a means to make explicit the assumptions that are often made concerning the location preferences of prehistoric people. A predictive model should be based on a theory of site location preferences, that can be quantified and tested against (unbiased) archaeological data sets (see also Whitley, in press). It is clear that the cultural component of these theories is at the moment virtually absent in predictive modelling practice. This paper intends to show that it is not impossible to include these variables into predictive modelling, and this will hopefully lead to further research into this subject.

2. Predictive Modelling and Environmental Determinism

The practice of predictive modelling for AHM is, at the moment, environmental deterministic in outlook and design. The predominant use of environmental input variables as archaeological site predictors, such as soil type, groundwater table, distance to open water and the like, has however been criticized on a number of occasions in academic literature (e.g. Wheatley 1993, 1996a, 2003; Gaffney and van Leusen 1995). The problems associated with environmentally based predictive modelling (van Leusen et al. 2002) can be summarized as follows:

- Archaeological theorists reject an understanding of past human behaviour in purely ecological/economical terms, and argue that social and cognitive factors determine this behaviour to a large extent, and should therefore be additional predictors for the presence and nature of archaeological remains;
- The maximum gain (a measurement of the degree of effectiveness of the predictive archaeological model over a 'by chance' model) of current predictive models seems to be about 70% (Ebert 2000, Wheatley 2003), which implies that a significant proportion of archaeological site locations cannot be predicted using purely environmental datasets; therefore, models based on environmental factors alone cannot be adequate tools for the prediction of archaeological site location.
- Unfortunately, social and cognitive factors seem to be difficult to model, and have so far only be studied for a very limited range of questions, based on very specialised data sets (mostly relating to the ritual prehistoric landscapes of Wessex in England e.g. Wheatley 1995; 1996b).

The American archaeologist Timothy Kohler observed this as early as 1988. "Why are the social, political, and even cognitive/religious factors that virtually all archaeologists recognize as factors affecting site location and function usually ignored in predictive modelling?" (Kohler 1988: 19). He gives the answer a few pages later: "Given the subtleties and especially the fluidity of the socio-political environment, is it any wonder that archaeologists have chosen to concentrate on those relatively stable, "distorting" factors of the natural environment for locational prediction?" (Kohler 1988: 21).

In essence, the situation has not changed since Kohler made these remarks. The present practice of predictive modelling is still very much environmentally deterministic. Cultural variables are not included in the models, resulting in predictions ultimately based on physical properties of the current landscape.

Practitioners of 'traditional' predictive modelling have mostly resisted the conclusion that their models will not be adequate because they lack the input of non-environmental data (e.g. Kvamme 1997). It is not because they do not want to include non-environmental factors; the problem is that these variables are regarded as being too abstract and intangible for use in a predictive model. Such models, so the argument goes, will therefore not become any better by investing valuable research time in mapping cultural variables. Several publications have focused on this apparent impossibility to incorporate non-environmental variables in predictive modelling (Wheatley 1996a; Stančič and Kvamme 1999 and Lock 2000). As a consequence, very few studies are available where an attempt is made to improve the gain of a model by incorporating non-environmental factors. As a consequence, the effect of including cultural variables into predictive models can at the moment not be assessed. The current situation is therefore characterized by a fundamental criticism of the environmental deterministic approach, coupled to a very poor insight into the potential of using cultural variables in predictive modelling.

Ultimately, the theoretical basis needed for the development of culturally based predictive models seems to be underdeveloped. It is evident that many models of prehistoric land use have been proposed for local case studies, but they are usually not generalized for application in a predictive modelling context, and often have never been tested in a rigorous way. A typical example of this is found in the theories regarding the location of Linear Band Ceramic settlements, in which a strong cultural component is supposed to be present (see Gaffney and van Leusen 1995), yet no predictive model based on this assumption has ever been made.

In conclusion, it may be suspected that the lack of progress in incorporating cultural variables into predictive modelling has less to do with the variables themselves, than with the geographic and interpretative models needed to operationalize them for predictive modelling. Many applications that claim to be exponents of cognitive archaeology, often framed in post-processual rhetoric, rely on the same techniques that are used for old-fashioned, processual studies, up to the extent where they might even be called 'cognitive deterministic'.

3. Cultural Variables: What are They?

It is important to realize that, when we are speaking of cultural variables, we can think of two ways of obtaining them. The first one is to consider them as measurable attributes of the archaeological sample that are not related to an environmental factor. So, instead of measuring for each individual site its soil type, elevation, distance from water and so on, we need to ask which properties of the site itself can be measured. These include properties like site location, size, functional type and period of occupation. These variables are clearly the expression of forms of social behaviour, although the interpretation of the specific behaviour involved may be subject to discussion. For ease of reference, these variables will be denominated cultural variables *sensu stricto*. In themselves, these variables are not extremely difficult to obtain, but the problems of analysing and interpreting archaeological site databases are manifold and must be addressed before these properties can actually be used for predictive modelling.

The second approach to defining cultural variables is to identify features of the landscape itself that can be interpreted as having cultural significance, such as sacred springs. These can be referred as to as *cultural landscape variables*, and are not necessarily excluded from 'traditional' predictive modelling, but often are not recognized as constituting a 'cultural' variable. It can, in fact, be argued that all environmental variables have a cultural component, even though the emphasis in traditional predictive modelling is usually on subsistence economy rather than symbolic meanings.

In order to make further use of cultural variables in predictive modelling, it is necessary to transform these variables into continuous variables: for each single variable a value should be available at any location within the study area. This is generally not a problem when using environmental data sets like soil maps or digital elevation models. Archaeological sites however are mostly represented as points, or in some cases as areas of a very limited extent. Similarly, landscape features that are considered to have cultural significance are in practice often also regarded as point-like, or at best linear in nature. A transformation is therefore necessary to use point-like or linear objects for predictive modelling. Two types of GIS techniques are currently available to perform this transformation: distance zonation and line-of sight analysis.

Distance zonation is customarily performed in environmental predictive modelling to obtain continuous variables from environmental features that are either linear (like rivers or coastlines) or point-like (springs).

In some cases, cost surfaces (also known as friction surfaces or effort models) are calculated by assigning a weight to landscape features according to their supposed accessibility. This technique is applicable to environmental as well as cultural variables. Distance decay models are used less often, and are based on demographic and/or political-economic models borrowed from human geography (e.g. Renfrew and Level 1979). These models are specifically relevant for cultural variables “, as they make it possible to incorporate the notion of interdependence of settlements (see e.g. Favory et al., 2003). A number of studies have appeared in recent years using line-of-sight analysis as a technique for obtaining continuous cultural variables, amongst others in attempts to demonstrate the ritual and symbolic meaning of the placement of monuments such as long barrows (Wheatley 1995; Gaffney et al. 1995). However, this type of analysis is certainly not restricted to cultural variables.

A good example of the use of cultural variables “ and distance zonation is provided by Ridges (in press), who attempted to include the distance to rock art sites in a predictive model in NW Queensland (Australia) – and actually succeeded in improving the gain of the model. This success is probably due to the fact that the ritual sites used are fixed in space, and can be mapped with relative ease in the specific environmental situation. The rock art sites are typical examples of what Whitley (2000) refers to as ‘fixed point attractors’. The precise moment of their creation may be unknown, but their position and symbolic meaning remain stable during a long period of time, making them long-term attractors for human activity.

In many other situations however, potential cultural variables are less stable, and cannot be mapped with ease. Examples of these include road networks, field systems, and the archaeological sites themselves, which all can have highly varying life-spans and may change in importance as attractors over time. In order to model the effects of long term land use development, it is necessary to use a technique that can deal with spatio-temporal variables, like dynamical systems modelling.

4. How to Proceed?

In order to remedy the current situation the following issues should be addressed:

- The identification of cultural variables that are significant for archaeological site location;
- The analysis of the utility of these variables for predictive modelling;
- The development and application of existing and new relevant modelling techniques; and
- The analysis of the performance of predictive models based on cultural variables compared to environmentally based models.

Following the recommendations in van Leusen et al. (2002), we suggest that four promising areas of research should be explored in order to improve on the current use of cultural variables in predictive modelling. These are:

1. A systematic analysis of the archaeological records and their aggregation into culturally meaningful entities

It is necessary to analyse what information can be extracted from existing archaeological databases that can be used in the definition of cultural variables. The aggregation of the archaeological contents of find spots into meaningful archaeological entities is currently not standardized. A possible solution could be to design an expert system that can be used for the classification of find spots. Apart from defining meaningful archaeological entities, the aggregation of multiple find spots into single archaeological sites is an important issue where the utility of the archaeological database for predictive modelling is concerned. Thirdly, a tendency can be observed recently to combine multiple archaeological sites into ensembles, which effectively constitutes a step away from the site level and towards a regional, landscape-based concept of archaeological entities. The main question here is: what types of aggregates can we distinguish, and can these be used as cultural variables “?

2. Analysis of the logistic position of settlements

It is anticipated that one of the most important cultural variables that can be used is the logistic position of the archaeological site itself. It has been shown by many researchers that the position of a settlement in a logistic network determines to a large degree its size and duration of occupation (e.g. Durand-Dastès et al. 1998). The development of techniques to analyse the logistic position of settlements can be addressed by looking at recent work in human geography.

3. The continuity of the cultural landscape

The cultural landscape has a historical dimension that strongly influences its use and usability. The existing cultural landscape influences the positioning of new sites. Kuna (1998), for example, mentions the importance of remnants of past landscapes on settlement location choice. Bell et al. (2002) demonstrated how later settlement in their Central Italian study area avoids areas settled in an earlier phase but conforms to paths from that earlier phase. Techniques to perform the long-term diachronical analysis needed for this type of modelling have been developed (e.g. by the Archaeomedes project; van der Leeuw 1998: Favory et al. 2003)

4. Line-of-sight analysis

In hilly areas and with certain site types that have a strong visual component (like burial mounds or megalithic tombs) line-of-sight analysis may be a type of analysis suitable for predictive modelling (see van Leusen 2002 chapters 6 and 16). The techniques for performing this type of analysis are well established. It will be noticed that the four research topics mentioned here all focus on cultural variables".

A thorough investigation of the use of cultural landscape variables would primarily involve the development of a decision rule framework that will incorporate the perception of the landscape into predictive modelling. In itself, this is an issue that merits attention, but the establishment of decision rules has always been at the heart of predictive modelling and is covered by a wide range of studies already. It would however be useful to start thinking about ways to model the perception of the landscape, as has been done by Whitley (2000), who tried to model the attractivity of the landscape for specific (economic) activities of Native American hunter-gatherers (see also Whitley, in press).

5. Conclusions

In a recent article on the use and abuse of statistical methods in archaeological site location modelling Woodman and Woodward (2002) come to the following conclusion: "There has been much criticism of locational studies since they are often based largely on environmental criteria. However, before researchers attempt to incorporate the more intangible social, cognitive, political and aesthetic factors, it would be wise to employ the appropriate statistical techniques required to deal with the complexities which already exist in even the most basic tangible and quantifiable environmental criteria".

Although we do not deny that many statistical problems still exist with regard to predictive modelling, we see no apparent reason why they should receive prime importance in further developing predictive modelling. In fact, the three main issues of statistical methodology, the development of adequate archaeological (and non-archaeological) data sets and the incorporation of non-environmental factors into the models are closely connected, and cannot be tackled in isolation. The papers presented in van Leusen and Kamermans (in press) show that new approaches to predictive modelling are starting to emerge, like exploring the potential of Bayesian statistical methods, using high resolution data for predictive modelling, and looking for ways to better embed predictive models into archaeological heritage management practice, for example by developing risk assessment methods. There is no doubt still a lot to do, and in this respect we have to disagree with Wheatley (2003) who argues that too much money is going into predictive modelling studies. He may be right that funding for GIS-related archaeological projects is mainly going into predictive modelling, but compared to the amount of money spent on all forms of prospection and excavation, investments made in predictive modelling seem relatively modest. Apart from that, investments for a thorough, scientific analysis of predictive modelling have been few and discontinuous.

We hope to have demonstrated that incorporating cultural variables into predictive modelling can be done, even though it is impossible to present a comprehensive overview in these few pages. It is up to the scientific community and public institutions to decide if this line of research is worth investing in. However, if the three issues mentioned above (statistical improvements, quality of the archaeological data set and the development of non-environmentally based models) are not tackled in the years to come, predictive modelling will remain to be criticized as a tool that is of dubious scientific quality, and not even capable of providing clear answers on where to spend money for archaeological research.

References

Bell, T., Wilson, A. and Wickham, A., 2002. Tracking the Samnites: landscape and communications routes in the Sangro Valley, Italy. *American Journal of Archaeology* 106 (2).

Durand-Dastès, F., Favory, F., Fiches, J.-L., Mathian, H., Pumain, D., Raynaud, C., Sanders, L. and van der Leeuw, S., 1998. *Des oppida aux métropoles. Archéologues et géographes en vallée du Rhône*. Anthropos, Paris.

Ebert, J. I., 2000. The State of the Art in "Inductive" Predictive Modeling: Seven Big Mistakes (and Lots of Smaller Ones), In Wescott, K. L. and Brandon, R. J. (eds), *Practical Applications of GIS For Archaeologists. A Predictive Modeling Kit*. London, Taylor & Francis. 129–134.

Favory, F., Fiches, J.-L. and van der Leeuw, S., 2003. *Archéologie et systèmes socio-environnementaux. Etudes multiscalaires sur la vallée du Rhône dans le programme ARCHAEOMEDES*. CRA-Monographies. Paris, CNRS Editions.

Gaffney, V. and van Leusen, M., 1995. Postscript – GIS, environmental determinism and archaeology: a parallel text, In Lock, G. and Stančič, Z. (eds), *Archaeology and Geographical Information Systems: A European Perspective*. London, Taylor & Francis. 367–382.

Gaffney, V., Stančič, Z. and Watson, H., 1995. The impact of GIS on archaeology: a personal perspective, In Lock, G. and Stančič, Z. (eds), *Archaeology and Geographical Information Systems: A European Perspective.* London, Taylor & Francis. 211–229.

Kohler, T. A., 1988. Predictive locational modelling: history and current practice, In Judge, W. L. and Sebastian, L. (eds), *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modeling*. Denver, US Bureau of Land Management. 19–59.

Kuna, M., 1998. The Memory of Landscapes, In Neustupný, E. (ed.), *Space in Prehistoric Bohemia.* Prague, *Academy of Science*. 77–83.

Kvamme, K. L., 1997. Ranters Corner: bringing the camps together: GIS and ED. *Archaeological Computing Newsletter* 47, 1–5.

Leeuw, S. E. van der (ed.), 1998. *The Archaeomedes Project: Understanding the natural and anthropogenic causes of soil degradation and desertification in the Mediterranean Basin. Research Results*. Office for Official Publications of the European Communities, Luxembourg.

Leusen, M. van, 2002. *Pattern to Process. Methodological investigations into the formation and interpretation of spatial patterns in archaeological landscapes*. PhD thesis.

Rijksuniversiteit Groningen, Groningen.

Leusen, M. van, Deeben, J., Hallewas, D., Zoetbrood, P., Kamermans, H. and Verhagen, Ph., 2002. *Predictive modelling for archaeological heritage management in the Netherlands. Baseline report for the BBO research program. Interim report*. Rijksuniversiteit Groningen, Groningen.

Leusen, M. van, and Kamermans, H. (eds), in press. *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*. NAR, ROB, Amersfoort.

Lock, G. (ed.), 2000. *Beyond the Map. Archaeology and Spatial Technologies*. NATO Science Series, Series A: Life Sciences – Vol. 321. IOS Press, Amsterdam.

Renfrew, C., and Level, E. V., 1979. Exploring dominance: predicting polities from centers, In Renfrew, C. and Cooke, K. L. (eds), *Transformations: Mathematical approaches to culture change*. New York, Academic Press. 145–167.

Ridges, M., in press. Understanding H-G behavioural variability using models of material culture: An example from Australia', In Mehrer, M. and Wescott, K. (eds), *GIS and Archaeological Predictive Modeling.* CRC Press, Boca Raton, Florida USA.

Stančič, Z. and Kvamme, K. L., 1999. Settlement Pattern Modelling through Boolean Overlays of Social and Environmental Variables, In Barceló, J. A., Briz, I. and Vila, A. (eds), *New Techniques for Old Times – CAA98. Computer Applications and Quantitative Methods in Archaeology*. BAR International Series 757. Oxford, Archaeopress. 231–237.

Verhagen, P., in press. Prospection strategies and archaeological predictive modelling. In van Leusen, M. and Kamermans, H. (eds), *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*. Amersfoort, NAR, ROB.

Wheatley, D. 1993, Going over old ground: GIS, archaeological theory and the act of perception. In Andresen, J., Madsen, T. and Scollar, I. (eds), *Computing the Past: Computer Applications and Quantitative Methods in Archaeology – CAA92*. Aarhus. 133–38.

Wheatley, D., 1995. Cumulative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological applications, In Lock, G. and Stančič, Z. (eds), *Archaeology and Geographical Information Systems: A European Perspective*. London, Taylor & Francis. 171–185.

Wheatley, D., 1996a. Between the lines: the role of GIS-based predictive modelling in the interpretation of extensive survey data, In Kamermans, H. and Fennema, K. (eds), *Interfacing the Past. Computer applications and quantitative methods in Archaeology CAA95*. Analecta Praehistorica Leidensia 28, 275–292.

Wheatley, D., 1996b. The use of GIS to understand regional variation in Neolithic Wessex. In Maschner, H. D. G. (ed.), *New methods, old problems: Geographic Information Systems in modern archaeological research*, Southern Illinois University at Carbondale Occasional Paper No. 23, (Carbondale, Illinois: Center for Archaeological Investigations, Southern Illinois University at Carbondale). 75–103

Wheatley, D. 2003, Making Space for an Archaeology of Place. *Internet Archaeology* 15. http://intarch.ac.uk/journal/issue15/wheatley_index.html

Whitley, T. G., 2000. *Dynamical Systems Modeling in Archaeology: A GIS Approach to Site Selection Processes in the Greater Yellowstone Region*. Unpublished PhD thesis. Department of Anthropology, Pittsburgh (PA), University of Pittsburgh.

Whitley, T. G., in press. A Brief Outline of Causality-Based Cognitive Archaeological Probabilistic Modeling. In van Leusen, M. and Kamermans, H. (eds), *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*. Amersfoort, NAR, ROB.

Woodman, P. E. and Woodward, M. 2002. The use and abuse of statistical methods in archaeological site location modelling. In Wheatley, D., Earl, G. and Poppy, S. (eds), *Contemporary Themes in Archaeological Computing*. Oxford, Oxbow Books.