PREDICTIVE MODELLING OF ROMAN SETTLEMENT IN THE MIDDLE TIBER VALLEY

1. Introduction

The aim of this research is to apply and evaluate the potential of predictive modelling as a means to further understanding ancient settlement in the middle Tiber valley. This area is one of the most intensively surveyed regions of the Mediterranean yet even here the extent of archaeological knowledge is uneven and some areas have yet to be subject to any systematic survey at all.

Since 1997, the area has been the focus of the British School at Rome's Tiber Valley Project, under the direction of Dr. Helen Patterson. The overall goal of the project is to study the changing landscape of the middle Tiber valley from protohistory to the medieval period (Fig. 1). It examines the impact of the growth, success and transformation of the city of Rome on the history of settlement and economy in the river valley (for project aims and results, see Patterson 2004; Patterson et al. 2004b; Coarelli, Patterson 2008).

2. The dataset

The data used in this analysis have been generated through a restudy of the material from John Ward-Perkins' South Etruria Survey carried out between the 1950s and 1970s, and their integration with data from more recent field work and numerous published surveys and excavations. The data were collected and integrated by two Leverhulme-funded fellows, Helga Di Giuseppe and Robert Witcher; the restudy of the South Etruria survey material was undertaken by twelve ceramic specialists (Patterson et al. 2004a, and papers in Coarelli, Patterson 2008). The data are housed in a relational database, which includes over 5500 findspots. As well as the management and archiving of data, this systemisation also opens up potential for spatial analyses (Kay, Witcher 2005; Witcher, Kay 2008). Each record includes spatial coordinates that enable visualisation within a GIS. The precision of field recording for the bulk of sites is 100 metres, and this relatively coarse resolution is reflected in the predictive models developed below.

3. Background of application

The process of data collation emphasised that, whilst most areas have produced some evidence of human activity, not all areas have been subject to
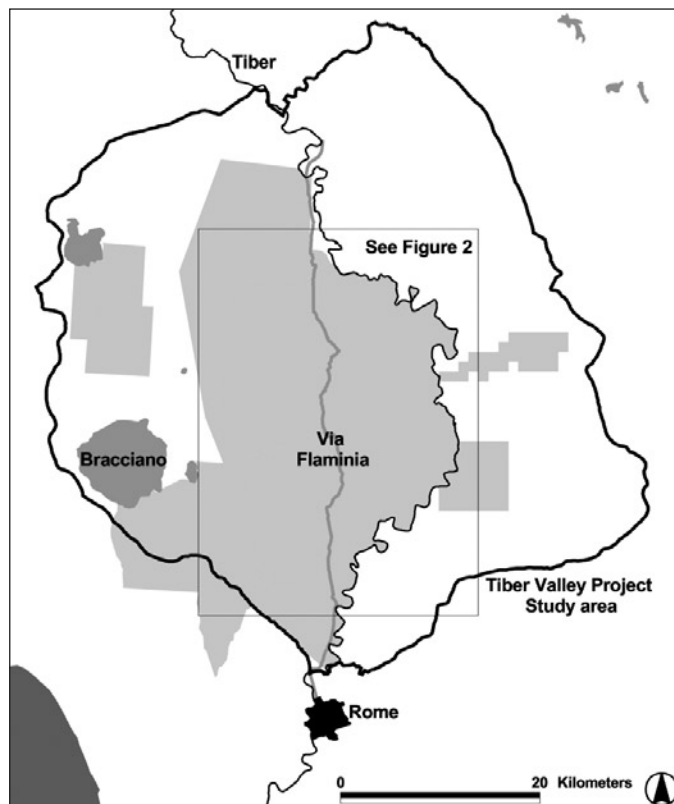
Fig. 1 – Location of Tiber Valley Project study area (black line), main field surveys (in grey) and predictive model case study area (boxed, see Fig. 2).

the same level of systematic study. For example, it became apparent that there was a marked contrast in the numbers and chronological development of sites on either side of the river Tiber (PATTERSON *et al*. 2004c). In order to assess whether this was a genuine pattern, or a product of uneven archaeological attention, a small field survey was undertaken in the Sabina Tiberina on the east bank (DI GIUSEPPE *et al*. 2002). The results suggest that the contrasting patterns of ancient settlement on either bank of the Tiber are likely to be a product of different histories of, and approaches to, landscape archaeology. In particular, the Corese survey was able to document higher densities of settlement than previously identified in the Sabina Tiberina, including sites from periods that have traditionally been difficult to recognise due to limited quantities of material culture (for example, the late antique).

Since the South Etruria survey commenced in the 1950s, some areas have been lost to quarrying and development, and many individual sites have been destroyed by erosion. It seems likely that many undiscovered sites have also been permanently lost. There are also several areas in South Etruria which were, and which remain, inaccessible, including military training areas and the Vatican Radio antenna farm. Predictive modelling offers the possibility to explore the potential archaeological significance of these different gaps and inconsistencies. Indeed, such work is of critical importance if we are to evaluate the significance of existing settlement patterns and trends. Furthermore, predictive modelling also has the potential to respond to a question of critical importance – what percentage of past settlement does field survey actually identify? (WITCHER in press).

4. Theoretical perspectives on predictive modelling

The majority of published applications of predictive modelling concern North American case studies (see papers in WESCOTT, BRANDON 2000; MEHRER, WESCOTT 2006; for brief introductions, see LOCK 2003, 168-70; CHAPMAN 2006, 157-159). In Europe, applications of the technique have concentrated on northern countries, particularly the Netherlands, Germany and France (LOCK, STANČIČ 1995; GARCÍA SANJUÁN, WHEATLEY 2002; VAN LEUSEN, KAMERMANS 2005; VERHAGEN 2007).

Some of this work is driven by specific research questions, but much concerns Cultural Resource Management where predictive modelling has come to form part of the planning process. There exist very few published applications of predictive modelling in southern Europe (see VAN LEUSEN 2002, 146-149); as well as GAFFNEY and STANČIČ's (1991) seminal study of the Croatian island of Hvar, there has also been work by KAMERMANS (2000) in the Agro Pontino to the south of Rome, by STANČIČ and VELJANOVSKI (2000) on the Croatian island of Brač, and by RUA (2009) in southern Portugal. Closely related GIS applications include work by VAN HOVE (2004) on "taskscapes" in Calabria and applications of Historic Landscape Characterization (HLC) in Greece and Turkey by CROW and TURNER (in press).

Of all the analytical techniques that have flourished since the widespread adoption and use of spatial technologies (for overview, see WHEATLEY, GILLINGS 2002), predictive modelling has been the most heavily critiqued, particularly from a European perspective. EBERT (2000), and more recently WHEATLEY (2004), have outlined a number of criticisms of inductive or data-driven predictive modelling. These criticisms can be grouped as two main points: 1) "prediction as explanation"; and 2) environmental determinism.

First is criticism of predictive modelling as explanation. However, predictive modelling does not aim to explain patterns, rather it aims simply to

identify patterns. It is the task of the archaeologist to explain and interpret those patterns, not the model. The second criticism is that predictive modelling is "anti-historical" because it «assumes [patterns are] wholly a product of the immediate surroundings of the individuals and communities» (WHEATLEY 2004). This argument relates to the observation that most variables used within predictive models are environmental (for example, slope, distance to water, etc.) and that social and cultural considerations are excluded. As a result, the approach is deemed to be "environmentally deterministic": constraints of the physical environment replace human agency. However, this wider argument is built on a confusion of correlation with causation. In common with any other statistical technique, it is inherent in predictive modelling that statistical association does not imply a causative relationship between variables. This returns to the first point: predictive modelling as descriptive not explanatory; it does not inherently exclude the agency of knowledgeable individuals or societies to structure their lives. Indeed, WHEATLEY (2004) himself notes that «this is not to deny that correlative predictive models may be telling us something about the behaviour of people in the past». North American applications have produced relatively powerful models based on environmental variables; it would be wrong to dismiss this predictive power on the grounds of how the results are (mis)interpreted.

In general, these arguments betray a series of broader misplaced and latent concerns. Firstly, that any attempt to involve environmental variables in an archaeological study is "determinism" by another name. Secondly, that quantification and statistical analyses are reductionist and, thirdly that "prediction" is antithetical to free agency. Briefly, these may be rebutted with the following responses: discussion of environment does not presuppose determination (for much more subtle approaches, including environmental affordance, see INGOLD 2000); statistical analyses, if appropriately used, become rigorous, repeatable investigations to support – not replace – interpretation; and, thirdly, as such studies merely describe, it is the responsibility of the archaeologist, not the models, to ensure that individuals are granted appropriate agency and that societies are given cultural autonomy. The best solution is to place less emphasis on "prediction" *per se*, and to foreground the approach as exploratory data analysis and pattern recognition, in other words, a form of data modelling and characterisation.

In the context of the current research, some other specific criticisms should also be addressed. Firstly, that predictive modelling is concerned only with sites and has failed to take broader theoretical developments about off-site activity into account. Whilst this is certainly a valid criticism, in relation to the present case study, the vast majority of data used here derive from site-based survey conducted fifty years ago, long before such theoretical and methodological developments. Whilst not ideal, it would be wrong to discard

these data as inadequate; indeed, as emphasised above, many of these sites no longer exist and thus these data form a unique documentary record which cannot be re-collected. A related issue is criticism that the technique deals only with a simple binary – site or no-site – and the possibility of more than one site per unit of land is not addressed (in the current study, South Etruria survey sites were recorded to the nearest 100 m, resulting in a spatial recording unit of 10000 m² which could theoretically contain more than a single site). Whilst recent high-intensity survey suggests that more than one site per 10000 m² is definitely a possibility in this area (Di Giuseppe *et al*. 2002), the precision of the original survey recording of 100 m precludes more detailed consideration at a higher resolution.

A further criticism has noted that many applications lack sufficient archaeological data and as a result fail to differentiate between sites of different dates and types. For example, in his study of the island of Brač, Stančič (2000) used 29 sites covering four centuries across 395 km². The present application uses 288 villas dating to the first century AD to train the model and then tests it against a reserve of a further 288 first century villas across a total area of 1100 km².

A final issue concerns the anachronistic nature of much environmental data used. It is important to include variables that are as chronologically relevant as possible; hence in the present application, modern land use was excluded from the final and most powerful model. However, such data should not be dismissed entirely. Whilst they may not necessarily inform about past settlement decisions, they may well shed light on recovery processes (in particular, visibility, see Terrenato, Ammerman 1996). Indeed, predictive modelling can be seen as an heuristic tool for exploring data and identifying influencing factors (such as post-depositional or recovery bias) that present significant problems for field surveyors.

5. Methodology

As described above, the site data are stored in a relational database and are linked, via SQL commands, with ESRI ArcView; the predictive models are developed using the Arc-WofE extension (data have been continuously migrated since the projects inception and are now analysed through ESRI ArcGIS; future analysis will make use of ArcSDM3.1, Sawatzky *et al*. 2004). The Weights of Evidence (WofE) methodology is part of a larger group of multi-criteria decision-making techniques and is commonly used, for example, for prospective mapping of mineral deposits. Goodchild (2007, 2008) provides an archaeological adaptation of multi-criteria decision-making which, rather than predict site locations, seeks to model agricultural productivity and population in the middle Tiber valley.

The WofE model uses statistical associations between known sites called training points (in this case, early imperial villas) and different map themes (such as geology, aspect and slope) in order to calculate a set of weights. It is therefore an inductive approach. These weights are then used to evaluate every possible combination of the different map layers in order to produce a single map (a unique conditions grid) showing probability of the presence of a site. The variables which the model identifies as important are considered for any significance in understanding ancient land use or perception of the landscape (for an archaeological application of WofE, see Hansen n.d.; for an alternative predictive methodology, also applied in the middle Tiber valley, see Espa *et al.* 2006).

The study area comprises *c.* 1100 km^2 in the middle Tiber valley, to the immediate north of the city of Rome (Fig. 1). A Digital Elevation Model was derived from contours and spot-heights from the Istituto Geografico Militare 1:25.000 map series. From this, maps of slope, aspect and topographical form (ridge, peak, valley, etc.) were derived. Other themes include geology, modern land use and rivers. Proximities to three "cultural" variables were also considered: Roman consular roads, contemporary Roman towns and the city of Rome.

For each theme, its relative weight (or influence) on site location is calculated by the WofE extension, by taking into account the spatial extent of each theme's attributes (for example, categories of geology) and the number of training sites present on each. Statistics evaluate the significance of the association between these sites and each theme and its individual attributes. For example, for geology, there is a strong aversion to alluvial areas; for topography there is strong preference for ridges and other convex topographical forms such as hilltops; an aversion to areas less than 100 m from watercourses; and an aversion to slope greater than fifteen degrees. The individual attributes of different themes can be categorised into varying numbers of classes (for example, four or eight classes of aspect) to assess the effects of this generalization on statistical association; similarly, various filters were used to derive topographical features generalized to different spatial scales (generally, see Wood 1996-2008). In each case, the classification producing the strongest association was used. Statistics were also calculated for themes based on proximity or distance. As the size and number of classes affects the output, these had to be carefully defined. Weights were graphed to identify significant cut off points. In the case of rivers, a simple binary theme of <100 m and >100 m to nearest watercourse was used; for proximity to roads, three categories were used (<1 km, 1-3 km, >3 km); for proximity to towns, three categories were used (<5 km, 5-10 km, >10 km); and for proximity to Rome, four bands were used (<20 km, 20-30 km, 30-40 km, >40 km).

On the basis of these statistics, different combinations of themes were used to generate unique combination grids, or probability surfaces (Kay,

WITCHER 2005, tav. 1). The model was developed as an exploratory process, with various themes introduced and excluded from the model in order to identify those combinations which were most predictive. The WofE extension includes a number of tests to ensure that the statistical assumptions of the model are not violated; in all the examples described here, these assumptions were upheld.

6. Results

The best (i.e. most powerful) model achieved with the environmental and cultural coverages listed above utilised just three environmental themes: geology (ten classes), topography (six classes) and slope (three classes). Particularly high probability combinations were level to gently sloping ground located on the tops of *tufo* ridges and spurs. The resulting probability surface is illustrated in Fig. 2. Darker shades indicate higher probability (for example, the narrow, fertile ridges in areas A and B); lighter shades indicate lower probability, most notably, the alluvial soils of the Tiber floodplain show up as pale areas.

The predictive model produced a highly complex and fragmented mosaic, in which areas of very low probability sit next to areas of high probability. These results may suggest the very careful localised positioning of sites in relation, particularly, to topographical form. The significant influence of the strongly dissected topography is reflected in the thin, parallel strips of land alternating between high and low probability.

Fig. 3 shows the cumulative percentage of background cells (i.e. random) and site cells (i.e. the reserve villa sites) against the posterior probability value. This demonstrates that both the random and the site groups comprise large numbers of low probability cells and fewer of higher probability. However, the slower accumulation of site cells indicates that a greater percentage of sites occur on higher probability cells. Overall, the model has moderate predictive power, but clearly offers a better-than-chance method of predicting site location. It effectively predicts *c*. 20% of villas in just *c*. 6% of the area, weakening to *c*. 53% of villas in *c*. 26% of the area.

The environmental and cultural themes used in the model have possible significance in terms of both past human behaviour (for example, site location preference) and archaeological recovery (for example, visibility). For example, people may have avoided building sites in valley bottoms due to flooding and on steep slopes due to erosion. However there are also post-depositional possibilities: alluvium may cover sites located in valley bottoms, whilst the lack of cultivation on steep slopes (there is little use of terracing in this area) means sites are less likely to be discovered if they did exist. It is, of course, not straightforward to distinguish between past settlement location
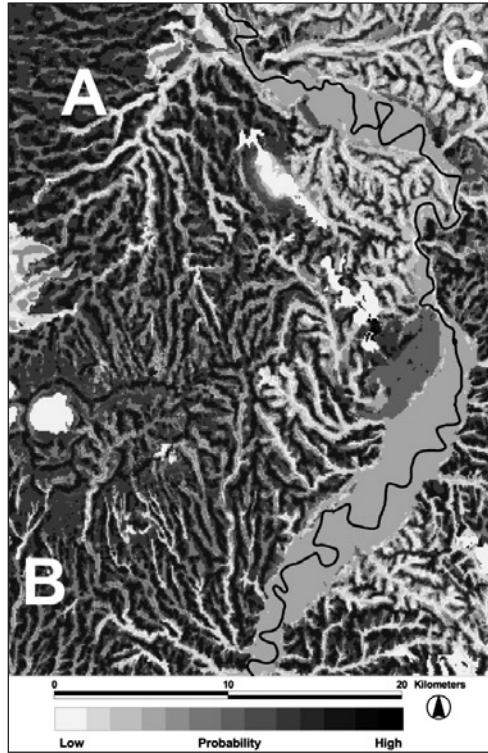
Fig. 2 – Predictive model for early imperial villas in the middle Tiber valley. For location, see Fig. 1. The river Tiber (in black) runs north-south. A = Central *Ager Faliscus*; B = area west of *Veii*; C = *Ager Foronovanus*.
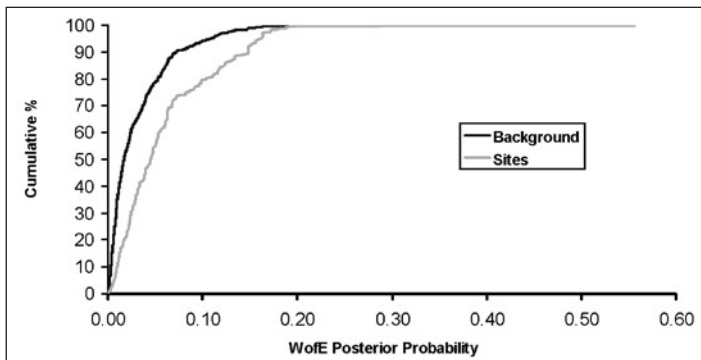

Fig. 3 – Comparison of cumulative percentages of background and sites cells against WofE posterior probability.

decisions and post-depositional and recovery issues. Of the themes used in the model, both cultural and environmental, it is clear that environmental themes have more predictive power. However, this is not to argue that they are more important in general than cultural factors, but that of the themes selected here, they have a more important role to play in prediction of where sites may be found today.

Land use and distance to Rome were employed in alternative models and both found to have high predictive power. However, both were excluded from further analysis as land use mapping refers to the modern landscape and seemed highly likely to reflect archaeological visibility (for example, pasture was low probability and vineyards high). Similarly, whilst it is possible that villa density was higher closer to Rome, archaeological activity/survey intensity has generally been much higher closer to the city. The current model therefore concentrates on those themes most clearly free of post-depositional and recovery problems. However, future work will attempt to use these and other themes specifically to distinguish between site location, post-deposition and recovery issues.

The model obviously predicts where sites might be found if the same survey methodology were employed again. As such, the model replicates existing biases (for example, surveyors may have concentrated their work in areas of high visibility where they believe they will achieve better results). Nonetheless, it still highlights potential unevenness in survey data. For example, the intensity of survey in the central *Ager Faliscus* (marked A on Fig. 2) and to the west of the city of *Veii* (B) is known to be particularly low. On the basis of their environmental similarity with other areas, the model suggests the probability of finding villas in these areas is high. However, these existing biases can be addressed through the integration of new and more systematic fieldwork into the model. Indeed, this addresses another of the criticisms made by WHEATLEY (2004), that predictive models are self-fulfilling prophesies as they reinforce existing biases – however, if modelling is treated as an iterative process, with new results added in, models can be constantly refined (RUA 2009). With the current model, it is noticeable that there is some difference in probabilities between the two banks of the Tiber in the top third of Fig. 2. Survey in both areas is relatively limited and the model is therefore more strongly influenced by discoveries in the southern half of the study area. However, the area on the west bank (B) in southern Etruria is of a similar nature in terms of topography, geology, land use, etc. to the better surveyed areas to its immediate south. In contrast, the area to the east of the river (C) is relatively unlike other surveyed areas of the Sabina Tiberina on the eastern bank of the Tiber. The addition to the model of results from recent survey work in this area (GABRIELLI *et al.* 2003; VERGA 2007) may improve confidence in the strength of this patterning.

Perhaps the most striking result is the similarity between early imperial villa location and some modern settlement as revealed by visualising 1:10.000 *Carta Tecnica Regionale* maps over the probability model (KAY, WITCHER 2005, tav. 2). The model, based on Roman villas, also distinguishes very precisely between nucleated medieval centres on the one hand (low probability), and their suburbs and sprawling discontinuous developments of the last thirty years on the other (high probability). Modern land use was explicitly excluded from the analysis and there is unlikely to be any correlation as a result of the preferential discovery of material during house construction, as the bulk of the data used pre-date such suburban developments. This might suggest possible similarities in the landscape perceptions, valuations or motivations influencing the settlement location decisions of both Roman villas and modern suburban settlement. Specifically, in contrast with Etruscan and Medieval settlement, these are open sites with little need or desire to nucleate. Most notably, they are on ridges. There are advantages to this, such as drainage and expansive vistas, but also disadvantages such as exposure to wind and inaccessibility of water. In the latter context, the widespread presence of Roman cisterns in this area is interesting. But probably most important is the fact that roads tend to follow the ridges. The *tufo* landscape across much of this area has created narrow ridges, divided by steep valleys. The consular road, the Via Flaminia, follows such a ridge and avoids the need to cross any river for more than thirty kilometres north of the Milvian Bridge outside Rome (Fig. 1). However, this is only the most impressive example of countless other ridge roads (that is, non-consular roads) which were excluded from this preliminary model because they are unevenly mapped and so may have skewed results.

Finally, the model of early imperial villa location was applied to samples of sites from other periods. Prediction of Etruscan and mid-republican sites was as efficient as the prediction of sites of imperial date on which the model was based. Further work is required to interpret this situation, but two (not necessarily mutually exclusive) explanations for this similarity in site location parameters can be postulated. First, that settlement in all three periods is similarly located (despite such significant events as the construction of consular roads and the economic pressures resulting from the emergence of the imperial metropolis of Rome). Second, that these patterns are largely the product of post-deposition and recovery, for example issues of visibility, which may effectively homogenise differences between the settlement preferences in different periods.

Assessment of the unevenness of the data is vital if other aspects are to be explored and developed. For example, work by GOODCHILD (2007, 2008; also GOODCHILD, WITCHER in press) on the agricultural production potential in the middle Tiber valley explores issues such as subsistence regimes, population and carrying capacities of land units through the use of historical and comparative evidence, as well as archaeological data. These issues can only

be reliably addressed through assessment of the completeness or otherwise of existing patterns. Predictive modelling is one method which may help improve understanding of the character of archaeological datasets.

7. Conclusion

In summary, the predictive power of the final model does not compare to that produced by some of the North American models. For example, Dalla Bona's (2000, 94) model of sites in northern Ontario located *c.* 84% of known sites within areas of high archaeological potential which comprised just *c.* 16% of the area. However, some key differences between North American and European models can be identified. Firstly, most US models deal with very few sites. It is possible that there was greater selectivity of prime landscape locations, whereas in a full and intensively exploited landscape such as that on the doorstep of imperial Rome, it would seem likely that choice was more about compromise. Further, in comparing the results presented here with North American models, it is apparent that in the hinterland of ancient Rome we are dealing with a more complex agricultural and territorial empire capable of significantly altering the environment and its potential to support large, sedentary populations (for example, cisterns for irrigated agriculture or inter-regional exchange). Pressures to supply the metropolis with agricultural goods and the desire or political need to live near Rome may have meant that environmental variables were increasingly less influential on settlement location over time.

Lower predictive power may be one reason for the lack of popularity of such work in the Mediterranean, but a more probable explanation is the very different developmental pressures and Cultural Resource Management processes. In particular, there are few areas (in Italy at least) about which absolutely nothing at all is known archaeologically; further, the archaeological record is densely-distributed; agricultural and developmental pressures are great, and the extent of areas to be assessed is comparatively small and can be subject to more intensive reconnaissance.

This study is a preliminary attempt to evaluate the possibilities of predictive modelling in the middle Tiber valley. In particular, it makes use of a generic modelling package; whilst this provides a useful initial framework, future work will seek to move away from a "blackbox" approach and seek to increase control over the process. Nonetheless, the initial results suggest that that predictive modelling is a useful heuristic tool to explore site location preferences and archaeological recovery issues in this part of the Mediterranean.

Stephen J. Kay
The British School at Rome

Robert E. Witcher
Department of Archaeology
Durham University

REFERENCES

Chapman H.P. 2006, *Landscape Archaeology and GIS*, Stroud, Tempus.

Coarelli F., Patterson H. (eds.) 2008, *Mercator Placidissimus. The Tiber Valley in Antiquity. New Research in the Upper and Middle River Valley*, Quaderni di Eutopia 8, Roma, Quasar.

Crow J., Turner S. in press, *Unlocking historic landscapes in the Eastern Mediterranean using historic landscape characterisation*, «Antiquity».

Dalla Bona L. 2000, *Protecting cultural resources through forest management planning in Ontario using archaeological predictive modelling*, in Wescott, Brandon 2000, 73-100.

Di Giuseppe H., Sansoni M., Williams J., Witcher R. 2002, *The Sabinensis Ager revisited: a field survey in the Sabina Tiberina*, «Papers of the British School at Rome», 70, 99-150.

Ebert J.I. 2000, *The state of the art in "inductive" predictive modelling: seven big mistakes (and lots of smaller ones)*, in Wescott, Brandon 2000, 129-134.

Espa G., Benedetti R., De Meo A., Ricci U., Espa S. 2006, *GIS based models and estimation methods for the probability of archaeological site location*, «Journal of Cultural Heritage», 7, 3, 147-155.

Gabrielli R., Guidi A., Santoro P. 2003, *Il progetto Galantina*, in J.R. Brandt, X. Dupré Raventós, G. Ghini (eds.), *Lazio e Sabina 1*, Roma, De Luca, 109-114.

Gaffney V., Stančič Z. 1991, *GIS Approaches to Regional Analysis: A Case Study of the Island of Hvar*, Ljubljana, Znanstveni Institut Filozofske Fakultete.

García Sanjuán L., Wheatley D.W. (eds.) 2002, *Mapping the Future of the Past. Managing the Spatial Dimension of the European Archaeological Resource*, Seville, Universidad de Sevilla.

Goodchild H. 2007, *Modelling Roman Agricultural Production in the Middle Tiber Valley, Central Italy*, University of Birmingham, Unpublished PhD thesis (http://etheses.bham.ac.uk/175/).

Goodchild H. 2008, *Modelling the productive potential of the middle Tiber valley*, in Coarelli, Patterson 2008, 769-796.

Goodchild H., Witcher R.E. in press, *Modelling the agricultural landscapes of Republican Italy*, in J. Carlsen, E. Lo Cascio (eds.), *Agricoltura e scambi nell'Italia tardo repubblicana*, Bari-Roma, Edipuglia.

Hansen D.T. 2004, *Describing GIS applications: Spatial statistics and weights of evidence extension in ArcView in the analysis of the archaeology sites in the landscape* (http://www.giscafe.com/technical_papers/Papers/paper054/, accessed 15/07/2004).

INGOLD T. 2000, *The Perception of the Environment. Essays on Livelihood, Dwelling and Skill*, London-New York, Routledge.

KAMERMANS H. 2000, *Land evaluation as predictive modelling: a deductive approach*, in LOCK 2000, 124-146.

KAY S., WITCHER R.E. 2005, *The Tiber Valley Project. The role of GIS and databases in field survey data integration and analysis*, «Archeologia e Calcolatori», 16, 113-137.

LOCK G. (ed.) 2000, *Beyond the Map. Archaeology and Spatial Technologies*, Amsterdam, IOS Press.

LOCK G. 2003, *Using Computers in Archaeology: Towards Virtual Pasts*, London, Routledge.

LOCK G., STANČIČ Z. (eds.) 1995, *Archaeology and Geographical Information Systems: A European Perspective*, London, Taylor and Francis.

MEHRER M.W., WESCOTT C.L. (eds.) 2006, *GIS and Archaeological Site Location Modeling*, London, Taylor and Francis.

PATTERSON H. (ed.) 2004, *Bridging the Tiber. Approaches to Regional Archaeology in the Middle Tiber Valley*, Archaeological Monographs of The British School at Rome 13, London, The British School at Rome.

PATTERSON H., BOUSQUET A., DI GIUSEPPE H., FELICI F., FONTANA S., WITCHER R., ZAMPINI S. 2004a, *Le produzioni ceramiche nella media Valle del Tevere tra l'età repubblicana e tardo antica*, «Rei Cretariae Romanae Fautores Acta», 38, 161-170.

PATTERSON H., DI GIUSEPPE H., WITCHER R.E. 2004b, *Three south Etrurian "crises": first results of the Tiber Valley Project*, «Papers of the British School at Rome», 72, 1-36.

PATTERSON H., DI GIUSEPPE H., WITCHER R.E. 2004c, *Progetto Tiber Valley. La Sabina e le due rive a confronto*, in A. GUIDI, S. PONCHIA (eds.), *Ricerche archeologiche in Italia e in Siria*, Padova, Sargon, 67-88.

RUA H. 2009, *Geographic Information Systems in archaeological analysis: a predictive model in the detection of rural Roman villae*, «Journal of Archaeological Science», 36, 2, 224-235.

SAWATZKY D.L., RAINES G.L., BONHAM-CARTER G.F., LOONEY C.G. 2004, *ArcSDM2: ArcMAP extension for spatial data modelling using Weights of Evidence, logistic regression, fuzzy logic and neural network analysis* (http://www.ige.unicamp.br/sdm).

STANČIČ Z., VELJANOVSKI T. 2000, *Understanding Roman settlement patterns through multi-variate statistics and predictive modelling*, in LOCK 2000, 147-156.

TERRENATO N., AMMERMAN A.J. 1996, *Visibility and site recovery in the Cecina Valley Survey, Italy*, «Journal of Field Archaeology», 23, 91-109.

VAN HOVE D. 2004, *Time and experience: taskscapes within GIS*, «Internet Archaeology», 16 (http://intarch.ac.uk/journal/issue16/vanhove_index.html, accessed 06/09/ 2007).

VAN LEUSEN P.M. *et al.* 2002, *Understanding digital archaeological landscapes*, in P. ATTEMA, G.J. BURGERS, E. VAN JOOLEN, M. VAN LEUSEN, B. MATER (eds.), *New Developments in Italian Landscape Archaeology*, BAR International Series 1091, Oxford, Archaeopress, 125-152.

VAN LEUSEN M., KAMERMANS H. (eds.) 2005, *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*, Nederlandse Archeologische Rapporten 29, Amersfoot, National Service for Archaeological Heritage.

VERGA F. 2007, *Ager Foronovanus I*, *Forma Italiae*, vol. 44, Firenze, Leo S. Olschki Editore.

VERHAGEN PH. (ed.) 2007, *Case Studies in Archaeological Predictive Modelling*, Archaeological Studies Leiden University 14, Amsterdam, Leiden University Press.

WESCOTT K.L., BRANDON R.J. 2000, *Practical Applications of GIS for Archaeologists: A Predictive Modelling Kit*, London, Taylor and Francis.

WHEATLEY D. 2004, *Making space for an archaeology of place*, «Internet Archaeology», 15 (http://intarch.ac.uk/journal/issue15/wheatley_index.html, accessed 15/07/2004).

Wheatley D., Gillings M. 2002, *Spatial Technology and Archaeology*, London, Taylor and Francis.

Witcher R.E. in press, *Missing persons? Models of Mediterranean regional survey and ancient populations*, in A. Bowman, A. Wilson (eds.), S*ettlement, Urbanisation and Population*, Oxford Studies in the Roman Economy 2, Oxford, Oxford University Press.

Witcher R.E., Kay S.J. 2008, *The Tiber Valley Project: the database and Geographical Information System*, in Coarelli, Patterson 2008, 417-423.

Wood J. 1996-2008, *Landserf* (http://www.landserf.org).

ABSTRACT

This paper discusses the results of an inductive predictive modelling experiment on Roman settlement data from the middle Tiber valley, Italy. The study forms part of the British School at Rome's Tiber Valley Project, which since its inception in 1997 has been assessing the changing landscapes of the Tiber Valley from protohistory through to the medieval period. The aim of this present study is to broaden understanding of settlement patterns via predictive modelling, and in particular to evaluate unevenness in field survey coverage, survey bias and past settlement location preferences. The predictive modelling method chosen was an application of the statistical Weights of Evidence extension for ESRI ArcView. The results highlight associations between Roman settlement and environmental themes that provide moderate predictive potential and suggest that further experimentation might prove valuable.