‘BELIEF’ in the past: Dempster-Shafer theory, GIS and archaeological predictive modelling

Shaun Canning

Abstract

This paper introduces the use of a new technique in archaeological predictive modelling which has particularly wide ranging appeal and application in the cultural heritage management sphere. Most predictive modelling programs are restricted to the use of certain types of data with a range of untested assumptions or caveats. Traditional probability based models are difficult to construct, and are often limited by the use of data sets which render the results statistically suspect. The method introduced here makes use of non-parametric statistical methods which are not hampered by imperfect raw data. In particular, this paper introduces the use of Dempster-Shafer theory in archaeological predictive modelling

Introduction

The majority of contemporary archaeological data in Australia are produced through cultural heritage management survey projects. These survey projects are conducted by contract archaeologists on behalf of exploration companies, developers or land managers. Academia has often criticised non-academic archaeologists for a lack of rigour, and the production of vast quantities of essentially useless archaeological survey data. Some of the perceived shortcomings of this corpus of data include poor sampling methods, insufficient survey coverage, non-standardised survey methods, visibility constraints and numerous other bias producing flaws or errors. Given the criticism aimed at the consulting archaeologists and the identified biases in the data, can the data produced by the consulting industry be put to good use? Or, as is currently the case, is this ‘grey literature’ simply dismissed as too problematic to be useful? This paper explores methods of utilising these data in predictive models, and discusses ways of allowing for the known biases, and incorporating this ‘error’ into the resultant models.

The study area and data

The project (Canning 2003) this paper describes was conducted in the Melbourne Metropolitan area between 1999 and 2003. One of projects main aims was to construct predictive models of Aboriginal archaeological site location within the relevant study area using the existing Aboriginal Affairs Victoria (AAV) sites database. The majority of the archaeological data contained in this database are the result of contract projects, and as such, contains many of the identified biases. Over 1000 sites and approximately 100 archaeological reports held by AAV were the source data for the predictive model developed later in this paper. The purpose of the predictive model is to provide local government planners with management tools that will enhance the identification, protection and management of Aboriginal archaeological sites within the relevant local government areas.

P.O. Box 21, Dampier, WA 6713, Australia.
Email: shaun@heritageservices.com.au.

Model building caveats

Models are essentially abstract representations of an observed or hypothesized phenomenon (Winterhalder 2000). The models developed here focus on very specific elements of Aboriginal behaviour in prehistory. They make use of ethnographic analogy, and utilise the theoretical perspectives of human ecology. It is specifically noted, however that the archaeological record cannot simply be read in terms of ethnoarchaeological understandings of past Aboriginal behaviour. Such an approach would ignore depositional and post-depositional processes, discard rates, erosion, and innumerable other factors which have created the archaeological record. All models ultimately provide nothing more than a generalised view of the archaeological record perceived by us through innumerable filters and obstructions. The methods presented later in this paper aim to extract meaning from data which may otherwise be of limited use in other forms of predictive modelling endeavours, and uses new techniques to achieve this end. These are therefore ‘self-consciously reductionist’ models’ (Winterhalder 2001: 14).

Predictive modelling

Predictive modelling in archaeology has its origins in the settlement pattern analyses of Julian Steward (1938) and Gordon Willey (1953). These pioneering studies focused primarily upon the relationship between regional environments and settlement patterns (Dalla Bona 1994; Kohler and Parker 1986). Out of the development of settlement pattern studies, and the increasing emphasis on scientific research methods, catchment analysis techniques were developed to investigate regional processes (Higgs and Vita-Finzi 1972; Vita-Finzi and Higgs 1970) that emphasise the relationships between people and their environment (Roper 1979). The ‘New Archaeology’ of the 1960s and the heightened interest in archaeological sampling techniques and data analysis methods (Binford 1964) led to a shift by some archaeologists away from ‘single-site’ archaeology to broader regional questions. The introduction of new technological tools (i.e. computers) has given practitioners the ability to interrogate greater quantities of data than had previously been possible (Dalla Bona 1994).

Amongst the first studies to explicitly state that a research goal was to predict actual archaeological site location was that of the South-western Anthropological Research Group (SARG) in the United States (Plog and Hill 1971). The SARG members reasoned that if the structure of a particular settlement system under consideration was known, then it should be possible, *a priori*, to predict the location of unknown archaeological sites (Plog and Hill 1971). Similar prediction-based research questions began to appear in the archaeological literature during the early 1970s (Green 1973; Thomas 1973, 1975). Although it has been claimed that the use of predictive modelling in pure research based activities has gradually declined (Dalla Bona 1994), the perceived benefits of predictive modelling in cultural heritage management (CHM) applications has fuelled the continued development of predictive modelling method and theory (Kohler 1988; Kohler and Parker 1986; McManamon 1984; Warren 1990).

Cultural heritage management agencies are generally mandated with the responsibility of managing and protecting a finite archaeological resource base for a relevant region, state or nation. This is attempted using various combinations of techniques, including excavation, site discovery and recording programmes, and more recently the application of predictive modelling. The responsibility to manage and protect the archaeological record means many agencies are constantly seeking improved methods to locate and record archaeological sites, particularly in the face of rapid development and urban expansion (MacNeill 1998). In this environment, the benefits of archaeological predictive modelling to cultural heritage managers appear obvious, even though no one is entirely certain as to how predictive modelling should be approached (Kohler 1988), and the approach eventually chosen is dependant on a great many variables. Paramount among these is determining the purpose for the model. Is the model to be a purely academic exercise, or is the model designed to 'red flag' (Altschul 1990) areas of archaeological sensitivity for planning and management agencies?

Types of predictive models

There are many types of predictive models in CHM and archaeology (Kohler 1988), which will be briefly summarised, together with their various theoretical perspectives, and the range of decisions involved (Gibbon 1998). Two major approaches are used in the construction of archaeological predictive models (inductive and deductive), particularly those developed within a CHM framework where their primary purpose is not explanation, but usually prediction. Explanation is perhaps the more desirable outcome of research activity, as theoretically grounded explanation is a more powerful tool than prediction alone.

Deductive Modelling involves deductive logic where the researcher moves from the abstract (theory) to the non-abstract (archaeological reality). Deductive predictive models commence from a certain theoretical perspective and proceed towards an understanding of extant archaeological data or phenomena primarily via explanation. Indeed, Kohler (1988:37) defines a deductive modelling approach as one that begins 'with a theory as to how people use a landscape and to deduce from that theory where archaeological materials should be located'. For example, a model of archaeological site location may be constructed using the theoretical perspectives of behavioural or human ecology (Butzer 1982). Once the model has been constructed, and middle-range theoretical issues such as discard rates, depositional and post-depositional processes have been incorporated, the modelling process can then turn to data gathering and the interpretation (explanation) of that data (Ebert 1988). Kohler and Parker (1986:432) state that a deductive model must:

(1) 'Consider how humans make choices concerning location. This requires considering a mechanism for decision-making, and an end for decision making; what is the goal?
(2) Specify the variables affecting location decisions for each significant chronological or functional subset of sites; and
(3) Be capable of operationalisation'.

In the realm of CHM predictive modelling, deductively based models are comparatively rare.

Inductive Modelling involves the researcher moving from detailed data to more generalised theory. In inductive predictive modelling, the researcher begins with site data and then makes estimates or inferences regarding the overall spatial distribution of archaeological material in that sampling universe (Kohler 1988; Neuman 1997). This type of modelling exercise makes use of existing data, such as site records held by management agencies, and is the most common approach to predictive modelling (Ebert 2000; Kohler 1988) particularly in CHM applications. This approach is also known as correlative or inferential modelling (Kohler 1988).

It has been argued that inductive predictive models are simply 'formal devices of pattern recognition' (Warren and Asch 2000:8). Inductive predictive modelling in CHM and archaeology is primarily concerned with attempting to find correlations between site location and the numerous environmental attributes considered relevant. These attributes are normally 'harvested' from existing contemporary geospatial data. Once (or if) any correlations are established, the process then moves on to 'modelling' the probable locations of further unknown sites (Ebert 2000). There are severe theoretical failings in the inductive modelling approach. Arguably, the most problematic issue with inductive or 'correlative' (Church et al. 2000) modelling is the almost implicit assumption that settlement decisions made by people in the remote past are somehow directly linked to geospatial attributes that can be derived or deduced from modern maps (Ebert 2000). While certain environmental attributes undoubtedly influenced prehistoric human settlement choices, not all of these attributes are static through time. Change in the environmental structure of a place through time is seldom taken into account in inductive models.

As an approach to predictive modelling, the inductively based model also has major appeal. The principal reason for the popularity of this approach is that the majority of the data required already exists in the form of site databases and geospatial map data. This significantly reduces the costs of any modelling project, which is another area of considerable appeal to the agencies that fund predictive modelling exercises (Church et al. 2000).

Mathematical vs. graphical modelling

In relation to the use of models, Rindos (1989:13) comments '…their potential weakness lies in their tendency to make us believe we have an insight into the data when we merely have created a mathematical epiphenomena'.

Dalla Bona (1994) discusses the second significant decision that is required when building a predictive model-the choice between a mathematical or graphical modelling methodology. Although the two may be combined, it has been the usual practice for modellers to choose one over the other.

Mathematical predictive models make use of any of the numerous multivariate statistical methods available in order to determine if correlations between archaeological site locations and the variables under analysis exist (Dalla Bona 1994). *Graphical* techniques often make use of modern computer hardware and software (particularly GIS) to develop a model as a series of map overlays of the relevant variables under consideration.

A third category of predictive model often seen in CHM literature and reports are known as intuitive models. Intuitive models are based upon a practitioners experience in the field, and their 'feel' for the archaeology of a particular area. Intuitive models are seldom tested (or are indeed testable) in an empirical sense. The 'predictions' will mostly be a series of statements such as 'sites will occur on terraces above waterways'. These are intuitive statements, and are not a 'predictive model' in the true sense of the term (Kohler 1988). This type of 'model' is common in Australian CHM, and many consulting reports (i.e. du Cros 1989, 1990, 1991) contain 'predictions' such as the preceding example. These types of intuitive statements are based upon the archaeologist's notion of where sites will be located (intuition), rather than any formalised research or sampling design (Altschul 1988; Moon 1993). These type of data are often referred to as 'expert knowledge'.

Attributes of predictive models

While predictive models may make use of a variety of methods and techniques, they can generally be divided between deductively or inductively based, and are either mathematical or graphical in design (this delineation is not absolute). All models however, regardless of method, should share a number of common attributes. Models should be testable, be simple enough to be useful, and must be able to be operationalised (Kohler 1988; Kohler and Parker 1986; Kvamme 1988a, 1988b; Moon 1993). Finally, because models are simple representations of reality, they are always fallible (Kohler 1988; Moon 1993). The choice between different approaches and styles of modelling depends as much upon the required outcomes, and the available data, as it does on the methods utilised. Either way, 'there is nothing inherently unscientific about either approach' (Warren 1990:90) and each method has its champions and its critics (Ebert 2000; Westcott and Kuiper 2000).

Predictive models, regardless of type, can be constructed to almost any scale. Contemporary predictive models of prehistoric archaeological site location have been constructed utilising a study area as small as a few hectares, up to very large undertakings such as the Minnesota Department of Transport's 'Mn/Model' (Brooks et al. 2000) which models prehistoric archaeological site location for the entire state of Minnesota (22,000,000 ha). Projects of this scale and budget are rare. The Minnesota Department of Transport (MN/Dot) spent $US4.5 Million on the project between 1995 and 2000, employing 49 people in various capacities. Although a rare level of commitment to one project, the Mn/Dot model has been shown to have saved the Minnesota Department of Transport $US 12 Million since 1998 (Minnesota Department of Transport 2000). The majority of predictive modelling projects however, are nowhere near this scale or scope.

The scope of predictive modelling projects is also broad. Models may be constructed to predict the location of archaeological sites from any temporal period or archaeological class. For example, predictive models have been developed in recent years to model the process of frontier settlement in the eastern United States (Zubrow 1990), to model the development of trade in the Great Lakes area of the United States (Allen 1990) and to address Palaeolithic and Mesolithic archaeological site location in the Southern Netherlands (Kamermans and Rensink 1999). The major limitations on predictive modelling are the availability of the appropriate data (in the case of inductive models) or the appropriate theoretical perspectives (in the case of deductive models) upon which to base the model(s). The scope of a model is also dependant upon the required outcomes of the modelling process. If the model is to predict archaeological site location on a contemporary landscape for management agencies to aid in the preservation of the archaeological record, then the model is constructed primarily for amenity, and not to answer specific research questions regarding human behaviour and archaeological processes in the past. However, overarching research paradigms are an absolute necessity as the ultimate goal of any CHM activity is the preservation of representative samples of complexly constructed archaeological landscapes. Research facilitates a better understanding of the nature of these complex archaeological landscapes.

Modelling sensitivity

It is generally considered impossible to construct a predictive model with the necessary spatial resolution (van Leusen 2002: 5-16) to predict the location of all individual hunter-gatherer-fisher archaeological sites within a region, particularly as the geographic scale of the model increases or the resolution of the spatial data decreases. Nor is the traditional reliance upon individual site based assessment well suited to the development of broad scale models of archaeological sensitivity or significance. This situation has led to the development of zone-based assessments of archaeological sensitivity and significance at landscape or regional scales (Altschul 1990; McConnell 1995, 2002; van Leusen 2002). Rather than attempting to predict the location or significance of individual sites, (or indeed the presence or absence of individual sites) zone-based assessments highlight those zones within a region that are expected to contain archaeological materials of various classes, and, *a priori*, significance. It must be remembered however, that it is the interpretation of archaeological material that determines significance, and not geomorphic or ecological predictions.

For instance, the deeply stratified alluvial deposits of the Maribyrnong Valley in the Australian state of Victoria (location of the Keilor and Green Gully sites) could reasonably be expected to contain buried prehistoric archaeological material. These comparatively rare occurrences are significant for their ability to illuminate past human behaviour in detail. However, it is impossible to predict the exact location of these rare phenomena with any chance of success. Knowing the geomorphic context in which these sites are likely to occur means the entire geomorphic unit can be regarded as archaeologically sensitive, and therefore likely to contain significant archaeological materials. This zone-based approach provides a superior method of identifying part or whole landscapes where archaeologically significant sites may be located rather than continuing to rely solely upon sporadic site surveys or test excavations.

Altschul (1990) utilised a zone-based methodology when developing the ‘red flag models’ of Mount Turnball in Arizona. Altschul (1990) viewed this approach as a more powerful tool to be used in everyday cultural heritage management contexts than individual site based assessments. Altschul’s (1990) methodology consisted of simply modelling three environmental variables (elevation, slope and aspect) believed to influence archaeological site location, and then plotting the relationship between these variables and the existing state archaeological database. The result of Altschul’s (1990) project was a series of ‘favourability’ maps which corresponded well to where archaeological sites were expected to occur, and did in fact occur. In terms of end usage, management agencies were handed a tool (favourability or sensitivity maps) that allowed them to ‘flag’ areas of greater sensitivity well in advance of any development related activities.

Anne McConnell (1995; 2002) developed similar zone-based models of archaeological sensitivity for management agencies in Tasmania and Victoria. McConnell (2002) assessed a series of environmental variables thought to have some bearing upon prehistoric Aboriginal archaeological site location. These attributes included distance to fresh water, slope, and access to flakeable stone. McConnell (2002), with the assistance of the Forest Modelling Branch of the Department of Natural Resources and Environment (DNRE), created a series of sensitivity maps using GIS that are to guide DNRE in the planning of logging and general management operations in Victorian forests. Josephine McDonald also utilised a similar method of sensitivity zoning in her study of a site on the Cumberland Plain near Sydney. McDonald (1996) based the sensitivity zones in her study primarily upon the level of previous ground disturbance that was observed throughout the study area. In this way, McDonald proposed that it was possible to identify entire landscapes that had undergone little disturbance since European settlement, and were thus more likely to contain undisturbed Aboriginal sites or ‘potential archaeological deposits’.

The major attraction of zone-based methods is that valid wide-area sensitivity models can be formulated (at a ‘macro scale’) where much of the data critical to statistically-based models is absent or cannot be determined (i.e. site absence vs. presence, statistically valid samples). For instance, the attempt by Lewis, MacNeill, and Rhoads (1996) to create a predictive model of archaeological site location in East Gippsland was not completely successful because of limitations in the available data sets and statistical complexity (McConnell 2002:29). Considering the frustrating array of limitations (i.e. visibility constraints) that can confront archaeological projects the sensitivity zone approach is arguably the most appropriate means of modelling the location of certain prehistoric archaeological materials. Rather than attempting to predict the locations of individual sites (as many models do), it is argued that a method of determining archaeological sensitivity based upon the relationship between existing known site data and key environmental attributes is a productive means of firstly identifying, and secondly preserving, archaeological material. In areas where it is possible to isolate particular geomorphic features that are known to be archaeologically sensitive, the zone-based approach is particularly useful. For example, the alluvial deposits of the Maribyrnong River valley are known to contain Pleistocene archaeological deposits of great scientific and cultural significance. Rather than attempting the impossible task of predicting where ‘individual’ buried sites lie, it is far simpler to zone this entire geomorphic context as ‘sensitive’ and impose limits to the type and extent of land altering development activities permitted within (or impacting upon) this zone. Within the various ‘zoned’ areas, it will still be necessary to conduct archaeological survey and research in order to update and improve the data for the model, as well as to ensure that any archaeological material located is recorded, and afforded the full protection of the relevant state CHM policies and legislation.

Establishing a predictive modelling program is not a particularly easy task. The researcher is faced with a myriad of choices, which have little bearing upon problems of an archaeological nature. The process is further complicated by the necessity of determining whether the model developed is to address management questions or provide insight into, or explanations of, human behaviour in the past. While both are equally valid pursuits, they are not always compatible aims. A management model that is designed to predict the most likely locations of prehistoric archaeological sites on a contemporary land surface is not seeking to explain why the archaeological material is located where it is. This type of model aims to identify where archaeological material is likely to be so as to avoid inadvertently destroying this material through development.

Case study

The model of archaeological sensitivity presented here was developed as part of doctoral research in the archaeology programme at La Trobe University (Canning 2003). This model is based upon the extensive collection of available archaeological data held by Aboriginal Affairs Victoria, and publicly available geospatial data sets. One map sheet has been chosen upon which to construct the model of archaeological sensitivity as it meets many of the criteria considered important for the modelling exercise. The VicMap 7822-1-3 1:25,000 Mapsheet was chosen for the modelling example because:

(1) It contains the important Keilor and Green Gully sites,
(2) It contains 276 registered Aboriginal sites,
(3) Approximately 10.5% of the total map sheet has been subject to archaeological survey, and
(4) It has been the subject of considerable previous archaeological research.

Summary of available data

The following environmental attributes were extracted by interrogating various spatial data sets using ArcView 3.2 to determine if relationships existed between archaeological site location and the variable in question. For many of the environmental variables analysed (i.e. soil drainage) it was simply impossible to determine any correlation between the archaeology and the variable (predominantly because of poor spatial resolution in the environmental data). The following section summarises the major environmental variables often believed to have considerable bearing upon archaeological site location.

Geology

The overall effect of geology on the distribution of archaeological sites in the area in question is difficult to determine. However, flakeable siliceous stone (i.e. silcrete) sources commonly occur at the junction of the basalt plains and the river valleys (Webb 1995) common in the area, and quartzite river cobbles are prolific in the various waterways.

Topography

A greater concentration of archaeological materials has been recorded at lower altitudes in the study area, and this is problematic in the construction of any sensitivity models. It is not clear if the lack of sites at higher altitudes accurately reflects prehistoric Aboriginal behaviour patterns, or is simply a product of bias in survey or the sites database. The effect of elevation on prehistoric Aboriginal site distribution is poorly understood, thus elevation is not a particularly strong predictive variable, particularly in areas that display relatively little topographic variation through large tracts of the subject area (as is the case here).

Distance to water

Distance to fresh water is the most often used environmental variable (van Leusen, 2002) in Australian hunter-gatherer archaeological modelling. Distance to water is used here in much the same manner as in any other project. The importance of access to potable water is considered one of the primary environmental factors affecting prehistoric land use decisions. Nearly two thirds of the recorded AAV sites (n = 276) for the area in question are within 0-100 metres of a fresh water source, and approximately 80% are within 200 metres (see Fig. 1).

Slope

Slope is a direct function of the topography of a region. In the present study area, slope is a variable with little real 'predictive power'. While almost 90% of all sites within the study area are located on landforms where the slope is between 0º and 10º, over 90% of the study area exhibits a slope of between 0º and 10º. The effect of survey bias on the distribution of sites per slope 'class' is also uncertain. While it would seem likely that Aboriginal occupation areas would more frequently be located on landforms displaying limited slope, it is not possible to quantify the relationship further. The areas displaying the greatest slope throughout the study area feature the least number of recorded sites. Again, the effect of uneven or biased survey coverage is not known.

Aspect

There is no clear pattern in the site data to suggest that one 'aspect' was preferred over any other. However, factors such as cold air drainage may still be shown to affect the choice of occupation sites.

Previously surveyed areas

One problematic attribute of the various data sets is the relationship between areas previously surveyed and the apparent proximity of sites to fresh water. While the proximity to fresh water is an important factor in the location of prehistoric archaeological sites, the location and extent (availability) of this resource will have changed markedly through both time and space. Contemporary survey coverage has tended to concentrate on those areas in

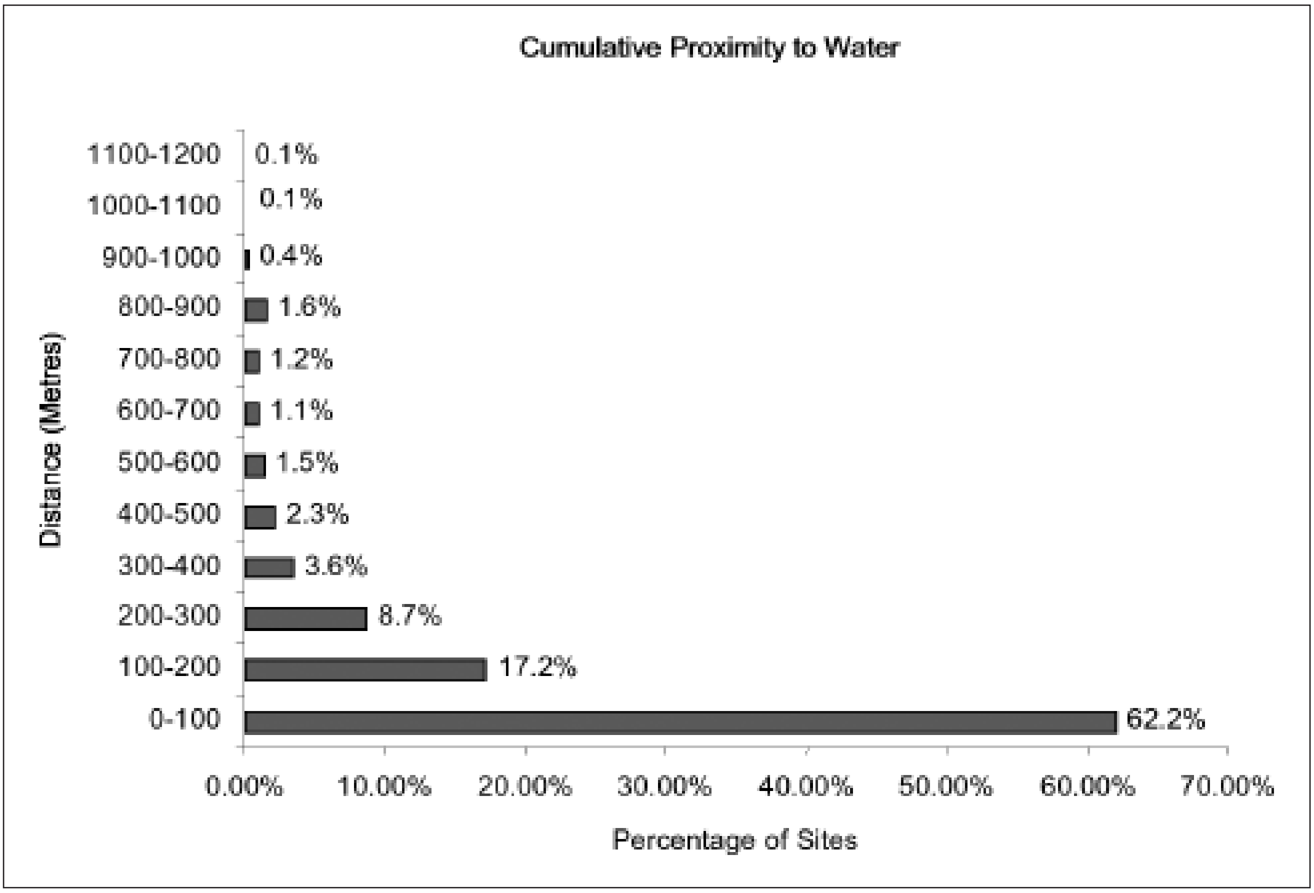

Figure 1 Distance to water for all known AAV sites types. The graph shows that 62.2% of all known AAV sites within the study area occur within 100 metres of a fresh water source. The 1:25,000 hydrology layer used in *ArcView* 3.2 for these calculations was modified to remove modern water features such as dams, reservoirs or drains.

close proximity to water, as most archaeologists 'know' that this is the area likely to yield the most sites. While this practice is common sense to a certain extent, it must also be remembered that a reliance on such 'expert' bias may result in an unrepresentative sample of the archaeological record. For instance, the AAV digital survey data of the study area were used to determine that a large proportion of survey activity has been undertaken within 200 metres of a source of fresh water (approximately 69%). Therefore, the observed site patterning may be a product of archaeological survey method as much as the result of prehistoric Aboriginal behaviour (Witter 1992:270).

'Weight of evidence' and Dempster-Shafer models

Management is essentially about an organizational response to uncertainty and risk. If all the parameters, choices and decisions of an organizations activity were known then active management would be redundant. In this regard, the management of archaeological material shares the same uncertainty and risk vocabulary as all other forms of resource management. Management uncertainty is 'inevitable in the decision making process' (Eastman 2001:23) and cultural heritage management operates within boundaries of considerable uncertainty. Uncertainty in archaeology can come from many sources. This paper considers several sources of uncertainty – namely uncertainty in the existing body of knowledge (i.e. no formal sampling, inconsistent survey intensity, overall lack of survey coverage, poor visibility) and uncertainty as to where other resources (sites or non-sites) are likely to be located. The ignorance of where undiscovered sites or non-sites are located introduces the risk that any existing but undiscovered archaeological resources may be destroyed through management processes that allow inappropriate activities to take place (i.e. unplanned development).

Biases in various data sources make it inappropriate to apply or utilise the wide range of parametric statistical techniques that are available in other archaeological pursuits (Orton 2000; Shennan 1997). This means that we cannot formulate answers to the predictive modelling questions in terms of binary opposites (yes/no - site/non-site) or standard probabilities. However, the masses of available data can be combined in a manner that produces valid results for any given data. Essentially 'weight of evidence' techniques are a means of combining various forms of evidence to support a hypothesis or hypotheses. These forms of evidence may be binary (i.e. presence or absence of sites) or may introduce other non-binary variables, which can be difficult to assimilate into models because the values are not binary (i.e. distance to water). For the purposes of the following discussion, sites are defined as geographic cells within a GIS that are known to contain archaeological materials and cover an area of 100 m^2 (10 m x 10 m). This is essentially the same definition of an archaeological site applied by AAV. Non-sites are the opposite of this, i.e. cells of 100 m^2 where no archaeological material is believed to exist. There are approximately 1.5 million cells of this size (100 m^2) in a map sheet such as the 7822-1-3-map sheet introduced previously.

Given the body of knowledge for the study area (the 'expert' knowledge and existing data) it is possible to begin to build a series of GIS layers that can be combined using various processes to produce a site 'likelihood' surface. A likelihood surface is not a quantitative probability statement. It does not state that a site will or will not exist at a specific point in space with a mathematical degree of precision. A likelihood surface is an indication that, on the balance of all the available evidence, a site is likely or unlikely to exist at that point in space. This type of analysis is particularly suited to cultural heritage management where so many of the parameters are either impossible to define, or where previous models are based upon untested hypotheses. The weight of evidence approach allows for the use of existing evidence in a manner that utilises particular aspects of Bayesian probability theory.

The GIS layers constructed here are based upon the enormous quantities of data generated by consultants and academics in the study area over the last 25 years. However, statements such as 'sites will occur on prominences in the landscape overlooking waterlines' are not easily converted to Boolean statements or queries for analysis in GIS. This is where the use of the raster GIS *IDRISI32* and its '*BELIEF*' module becomes indispensable. ArcView 3.2 was used for the majority of this project. While this is an adequate piece of software in its own right, *IDRISI32* offers a suite of powerful tools based upon Dempster-Shafer belief theory, which is an extension of Bayesian probability theory. 'The basic assumptions of Dempster-Shafer theory are that ignorance exists in the body of knowledge, and that belief for a hypothesis is not necessarily the complement of belief for its negation' (Eastman 2001:34). The workings of the *IDRISI32* '*BELIEF*' module are largely beyond the scope of this paper, however the 'Dempster-Shafer rule of combination provides an important approach to aggregating indirect evidence and incomplete information' (Eastman 2001: 36) in GIS-based modelling.

In order to model a likelihood surface we need to decide what is being hypothesized. In this case, the relatively straightforward binary opposites ('site' and 'non-site') are the two basic elements (hypotheses) of the decision frame. Evidence to support one or other is proffered from numerous sources. In this case, the evidentiary layers are distance to water, slope, and proximity to known sites. None of these attributes is easily described by internal binary relationships (i.e. they are not interval measurements, but are more like ratio measurements). For instance, the statement 'sites will occur at between 0 and 200 metres from a source of potable water' cannot easily be transformed into statements (map algebra) understood by GIS. Prior knowledge and experience would suggest that this is a valid statement for much of the archaeology of Australia, but this does not allow us to determine the relationship between distance to water and sites (i.e. are more sites really located closer to water than further away?).

The '*BELIEF*' module in *IDRISI32* contains numerous procedures that allow for the variable nature of the model attributes to be accounted for. When these processes are used on a 'distance to water' layer for example, the '*BELIEF*' module can be programmed to take into account that the further away from a source of potable water we move the more likely it is that each cell will be a non-site. From the available sites data for the study area, we know that approximately 80% of all known sites occur within 200 metres of a source of potable water (inclusive of biases). The falloff in site frequency at distances greater than 200 metres from water is shown in Figure 1. Site frequencies

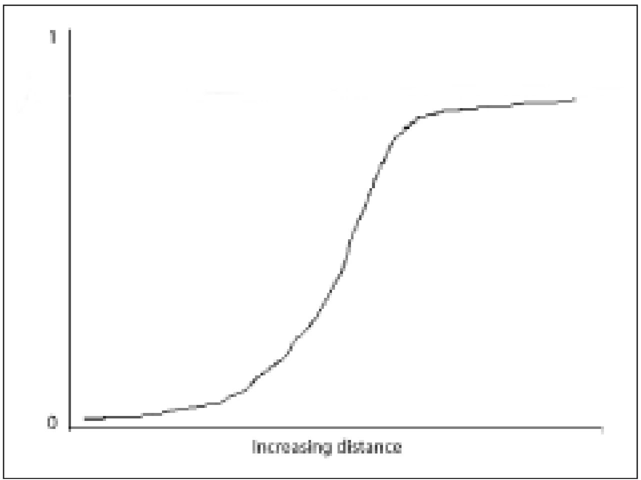

Figure 2 Graph showing the sigmoidal nature of the relationship between archaeological sites and water sources as the distance to water increases (0 = site 1 = non-site).

decline at distances greater than 200 metres from potable water, reaching almost zero beyond 1000 metres. The relationship of site proximity to water is shown as a sigmoidal (s-shaped) curve (Fig. 2).

Figure 2 shows the manner in which distance decay can be viewed graphically. This type of curve best represents the relationship between the distance to water and the number of sites, as there are no 'hard' boundaries delineating where site distributions and densities change or do not change. Close to water sources, the probability of encountering a non-site is low (i.e. nearer 0). As we move further away from a source of water, the probability of encountering a non-site increases to the point where it is theoretically 100%. A sigmoidal curve demonstrates this cumulative nature of distance to water and site distribution. As we move further away from the water source, the closer we are to the *theoretical* point at which no further sites will be found (i.e. the likelihood of a non-site approaches 100%). Naturally, we also move closer to encountering the next source of potable water, so there are limits to the application of this technique.

Other landscape attributes may be modelled in a similar manner. Slope is the other variable for which we have a significant amount of prior or existing expert knowledge, as well as the limited (and biased) quantitative data from GIS analyses. The accumulated data suggests that sites will occur in areas where the slope is between 0º and approximately 10º and that the sites will most commonly occur near an area of topographic change (i.e. where the plains meet the hill slopes of the river valleys). The same GIS processes described above can be run on these attributes to create a series of 'likelihood' surfaces to be incorporated into the final weight of evidence model. The known data for the relationship between site location and slope, for instance, can be processed to create two separate surfaces that show the likelihood of the occurrence of both sites and non-sites.

Because there is a degree of uncertainty in the data, and the completed modelling exercise should reflect this, the layers must be 'scaled' or weighted to ensure that the results do not indicate 100% certainty for any predicted value. *IDRISI32* makes this process comparatively easy. Layers can be scaled (i.e. multiplied) by any factor to reflect the degree of uncertainty. For instance, the Distance to Water layer used in the modelling exercise here has been weighted using a factor of 0.8 (80%). This simply means that the known distribution of archaeological sites (i.e. approximately 80% within 200 metres of water) has been accounted for, while factoring in an estimate of the uncertainty (i.e. the other 20% that occur at varying distances greater than 200 metres from water). *'FUZZY'* logic is applied within *IDRISI32* to model those cells where it is unlikely that a site will occur (non-site). Table 1 presents a summary of the layers that were created in *IDRISI32* for incorporation into the final aggregated *'BELIEF'* model. A comprehensive discussion of the operation of the *IDRISI32 'BELIEF'* and *'FUZZY'* functions is provided by Eastman (2001).

When the various layers are entered into the *IDRISI32 'BELIEF'* module, the surface produced shows the likelihood of a cell being a non-site. Because uncertainty has been factored into this model, no values greater than 0.8 are used. Where a value of 0.8 is shown, the model predicts that there is a 80% likelihood that the cell in question will be a non-site. Where the value returned by the model is low, i.e. 20%, the model predicts that there is an 80% likelihood that the cell is a site (i.e. 20% likelihood of non-site equals an 80% likelihood of a site).

Interpreting the model

The site likelihood surface generated from the available data should not be seen as a definitive probabilistic model. Interpreting this likelihood surface is relatively straightforward. Where the resultant value for any cell is high (i.e. >0.70) there is a high likelihood of encountering cells (remembering that each cell represents 100m^2) that do not contain any archaeological sites (i.e. non-sites). Where the value is low (i.e. <0.20) there is a high likelihood of encountering cells that will contain archaeological material.

It is explicitly recognised that other environmental and socio-cultural attributes will affect the presence or absence of archaeological sites in any area. In the current study area, for example, the location and distribution of siliceous lithic material can be used as a predictor variable for locating prehistoric silcrete quarries. The difficulty however, is that this material outcrops within the river valley slopes already known to be archaeologically significant (Webb 1995). The river valleys are already given significant 'weight' within the model due to the inclusion of 'slope' as a variable, so no extra 'weight' was thought to be required.

Conclusion

While it is undeniable that statistically rigouress quantitative predictive models should be the ultimate aim of any research driven modelling exercise, the reality is that this is often an impossible goal. In the case presented here, the majority of the data available for use already existed in the AAV database. However, these data are not suited for use in traditional statistical analyses, as various biases have been introduced during their collection (i.e. no standardised survey methods, poor sampling techniques, no visibility quantification). Rather than ignore this huge data set, the methods presented here specifically acknowledge that the data are biased, and allow them to be incorporated into a model of archaeological sensitivity. Nor are models of this type intended to supplant archaeological survey. The intention of the model presented here is to focus management attention on those areas believed to be the most archaeologically sensitive. The resultant 'likelihood' surface is an ideal tool for use by local government planners and other statutory authorities to flag those areas that must have archaeological investigations undertaken prior to any land altering activities. The models are not intended to explain prehistoric human behaviour of the area in question, but to further assist in the preservation of representative samples of the archaeological record for the researchers of tomorrow.

Acknowledgements

The Basalt Plains Archaeology Project was made possible through the generous support of the Australian Research Council via an APA (I) SPIRT grant incorporating a three-year PhD scholarship. Generous financial support was also forthcoming from the industry partner, Aboriginal Affairs Victoria. Dr Richard Cosgrove and Dr David Frankel of La Trobe University supervised the project, and their contribution is gratefully acknowledged. Thanks are due to the huge number of individuals who were involved with the project in one-way or another. Special thanks to Ken Mulvaney, Sean Ulm, Christine Martin, Brian Egloff and Harry Allen for comments on earlier drafts of this paper.

References

Allen, K. 1990 Modelling early historic trade in the Eastern Great Lakes using Geographic Information Systems. In K. Allen, S. Green and E. Zubrow (eds) *Interpreting Space: GIS and Archaeology*, pp. 319-329. London: Taylor and Francis.

Allen, J., Cosgrove, R. and Brown, S. 1988 New archaeological data from the Southern Forests Region, Tasmania: A preliminary statement. *Australian Archaeology* 27:75-88.

Altschul, J. 1988 Models and the modelling process. In J.W. Judge and L. Sebastian (eds) *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling*, pp. 61-96. Denver: U.S. Department of the Interior, Bureau of Land Management.

Altschul, J. H. 1990 Red Flag Models: The use of modelling in management contexts. In K. Allen, S. Green and E. Zubrow (eds) *Interpreting Space: GIS and Archaeology*, pp. 226-238. London: Taylor and Francis.

Binford, L. R. 1964 A consideration of archaeological research design. *American Antiquity* 29**:** 425-441.

Brooks, A., Hudak, G. J., Gibbon, G. and Hobbs, E. 2000 Management summary. In *Mn/Model: A Predictive Model of Pre-contact Archaeological Site Location for the State of Minnesota*, pp. 2-8. St Paul: Minnesota Department of Transport.

Butzer, K. 1982 *Archaeology as Human Ecology*. Cambridge: Cambridge University Press.

Canning, S. 2003 Site Unseen: Archaeology, Cultural Resource Management, Planning and Predictive Modelling in the Melbourne Metropolitan Area. Unpublished PhD thesis, Department of Archaeology, School of European and Historical Studies, La Trobe University. Bundoora, Melbourne.

Church, T., Brandon, R. J. and Burgett, G. R. 2000 GIS application in archaeology: Method in search of theory. In K.L. Wescott and R.J. Brandon (eds) *Practical Applications of GIS for Archaeologists: A Predictive Modelling* Kit, pp. 135-155. London: Taylor and Francis.

Dalla Bona, L. 1994 Cultural Heritage Resource Predictive Modelling Project. Volume 1: Introduction to the Research. Ministry of Natural Resources, Ontario.

du Cros, H. 1989 The Western Ring Road Archaeological Study. Stage 2, Volume 2. Report to VicRoads, Melbourne.

du Cros, H. 1990 The Sydenham Corridor: A Cultural Heritage Study. Victoria Archaeological Survey, Melbourne.

du Cros, H. 1991 The Werribee Corridor: An Archaeological Survey. Aboriginal Affairs Victoria. Melbourne.

Layer Name	Cell Hypothesis	Description	Justification(s)
Known Site	Site	Those cells where a known site exists, plus all cells within 300 m of a known site. FUZZY logic applied, using sigmoidal monotonically decreasing curve. The further away from a known site, the less likely it is that a cell will be a site.	Other sites will occur in close proximity to existing sites. As the distance between sites increases, so does the likelihood that a cell will be a non-site. The distribution of material from prior surveys demonstrates that the presence of sites in a cell is strongly influenced by the location of other archaeological material.
Distance to Water	Non-Site	Cells greater than 300 m to a source of potable freshwater. Cells between 0-300 m have FUZZY logic applied using a sigmoidal monotonically increasing curve. The greater the distance away from potable water, the higher the likelihood a cell is a non-site.	Distance to freshwater affects the distribution of site(s). The exact pattern is not known, although the overwhelming majority of sites in the study area (~90%) that occur within 300 m of a permanent water source. This layer is weighted to reflect this phenomenon.
Slope	Non-Site	Cells where the slope angle exceeds 25º are more likely to be non-sites. Those less than 25º are more likely to contain sites. FUZZY logic is applied, using a sigmoidal monotonically increasing curve. Those values between 0º and 25º are weighted more heavily than those greater than 25º.	The distribution of archaeological sites shows that sites tend to occur on slopes of between 0º and 25º. This is not to say that no sites will occur on slopes greater than 25º, rather that it is less and less likely as the slope increase. The FUZZY logic applied factors this into the aggregation of evidence.

Table 1 The various layers created for the 7822-1-3 Mapsheet, and the processes applied to them within IDRISI32.

Eastman, J. R. 2001 *IDRISI32 Release 2: Guide to GIS and Image Processing*. Worcester, Massachusetts: Clark Labs, Clark University.

Ebert, J. 1988 Modelling human systems and "Predicting" the archaeological record: The unavoidable relationship between theory and method. *American Archaeology* 7:3-8.

Ebert, J. 2000 The state of the art in 'Inductive' predictive modelling: Seven big mistakes (and lots of smaller ones). In K. L. Westcott and R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists: A Predictive Modelling Kit*, pp. 129-134. London: Taylor and Francis.

Gibbon, G. 1998 Archaeological predictive modelling: An overview. In *Mn/Model: A Predictive Model of Pre-contact Archaeological Site Location for the State of Minnesota*. St. Paul: Minnesota Department of Transport.

Green, E. L. 1973 Location analysis of prehistoric Maya sites in Northern British Honduras. *American Antiquity* 38:279-293.

Higgs, E. S. and Vita-Finzi, C. 1972 Prehistoric economies: A territorial approach. In E. S. Higgs (ed.) *Papers in Economic Prehistory*, pp. 27-36. Cambridge: Cambridge University Press.

Kamermans, H. and Rensink, E. 1999 GIS in Paleolithic archaeology: A case study from the Netherlands. In L. Dingwall (ed.) *Computer Applications and Quantitative Methods in Archaeology* Proceedings of the 25th Anniversary Conference, Archaeology in the Age of the Internet, University of Birmingham, April 1997. Oxford: Archeopress.

Kohler, T. A. 1988 Predictive locational modelling: History and current practice. In J. W. Judge and L. Sebastian (eds) *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive* Modelling, pp. 19-60. Denver: U.S. Department of the Interior, Bureau of Land Management.

Kohler, T. A. and Parker, S. C. 1986 Predictive models for archaeological resource location. In M. B. Schiffer (ed.) *Advances in Archaeological Method and Theory*, Volume 9, pp. 397-452. New York: Academic Press.

Kvamme, K. 1988a Development and testing of quantitative models. In J. W. Judge and L. Sebastian (eds) *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling*, pp. 325-428. Denver: U.S. Department of the Interior, Bureau of Land Management.

Kvamme, K. 1988b Using existing Archaeological survey data for model building. In J. W. Judge and L. Sebastian (eds) *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive* Modelling, pp. 301-323. Denver: U.S. Department of the Interior, Bureau of Land Management.

Lewis, A., MacNeill, R. and Rhoads, J. 1996 *East Gippsland Aboriginal Cultural Site Modelling*. Melbourne: Department of Conservation and Natural Resources, Aboriginal Affairs Victoria, and Centre for Resource and Environmental Studies, Australian National University.

MacNeill, R. 1998 GIS on the brink: GIS applications at Aboriginal Affairs Victoria. Paper presented at AURISA '98 - the 26th Annual Conference of AURISA. Perth, Western Australia, November 1998.

McConnell, A. 1995 Archaeology Potential Zoning: A Strategy for the Protection of Aboriginal Archaeological Sites in Tasmanian State Forests. Forestry Tasmania, Hobart.

McConnell, A. 2002 Aboriginal Heritage Management in Victorian Forests. Draft Project Report Volume 2. Central Highlands Region. Department of Natural Resources and Environment, Melbourne.

McDonald, J. 1996 The conservation of landscapes: A strategic approach to Cultural Heritage Management. In S. Ulm, I. Lilley and A. Ross (eds) *Australian Archaeology: Proceedings of the 1995 Australian Archaeological Association Conference*, pp. 113-121. St. Lucia: University of Queensland.

McManamon, F. P. 1984 Discovering sites unseen. In M.B. Schiffer (ed.) *Advances in Archaeological Method and Theory*, Volume 8, pp. 223-292. New York: Academic Press.

Minnesota Department of Transport 2000 History of the Mn/Model Project Minnesota Department of Transport, On-Line Report at http://www.mnmodel.dot.state.mn.us/pages/history.html.

Moon, H. 1993 *Archaeological Predictive Modelling: An Assessment*. Victoria, British Columbia: Resources Inventory Committee.

Neuman, W. L. 1997 *Social Research Methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon.

Orton, C. 2000 *Sampling in Archaeology*. Cambridge Manuals in Archaeology. Cambridge: Cambridge University Press.

Plog, F. and Hill, J. 1971 Explaining variability in the distribution of sites. In G.J. Gummerman (ed.) *The Distribution of Prehistoric Population Aggregates*. Prescott, Arizona: *Prescott College Anthropological Reports* 1:7-36.

Rindos, D. 1989 Diversity, variation and selection. In R.D. Leonard G.T. and Jones (eds) *Quantifying Diversity in* Archaeology, pp. 13-23. Cambridge: Cambridge University Press.

Roper, D. C. 1979 The method and theory of Site Catchment Analysis: A review. In M. B. Schiffer (ed.) *Advances in Archaeological Method and Theory*, Volume 12, pp.119-140. New York: Academic Press.

Shennan, S. 1997 *Quantifying Archaeology*. Edinburgh: Edinburgh University Press.

Steward, J. 1938 *Basin-Plateau Aboriginal Socio-Political Groups*. Washington, D.C.: *Bureau of American Ethnology Bulletin* 120.

Thomas, D. H. 1973 An empirical test for Steward's model of Great Basin settlement patterns. *American Antiquity* 38:155-176.

Thomas, D. H. 1975 Non-site sampling in archaeology: Up the creek without a site. In J.W. Mueller (ed.) *Sampling in* Archaeology, pp. 61-81. Tucson: University of Arizona Press.

van Leusen, M. 2002 Pattern to Process: Methodological Investigations into the Formation and Interpretation of Spatial Patterns in Archaeological Landscapes. Unpublished PhD thesis, Groningen Institute for Archaeology, University of Groningen, The Netherlands.

Vita-Finzi, C. and Higgs, E. S. 1970 Prehistoric economy in the Mount Carmel area of Palestine: Site Catchment Analysis. *Proceedings of the Prehistoric Society* 36:1-37.

Warren, R. 1990 Predictive modelling in archaeology: A primer. In K. M. Allen, S. W. Green and E. Zubrow (eds). *Interpreting Space: GIS and* Archaeology, pp. 90-111. London: Taylor and Francis.

Warren, R. E. and Asch, D. L. 2000 A predictive model of archaeological site location in the Eastern Prarie Peninsula. In K. L. Wescott and R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists: A Predictive Modelling Kit*, pp. 5-32. London: Taylor and Francis.

Webb, C. 1995 The Identification and Documentation of Silcrete Quarries. Aboriginal Affairs Victoria, Melbourne.

Westcott, K. L. and Kuiper, J. 2000 Using GIS to model prehistoric site distributions in the Upper Chesapeake Bay. In K. L. Wescott and R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists: A Predictive Modelling* Kit, pp. 59-72. London: Taylor and Francis.

Willey, G. 1953 *Prehistoric Settlement Patterns in the Viru Valley, Peru*. Washington, D.C.: *Bureau of American Ethnology Bulletin* 155.

Winterhalder, B. 2000 Models. In J.E. Terrell (ed.) *Darwin and Archaeology: A Handbook of Key* Concepts, pp. 201-223. Westport, Connecticut: Bergin and Garvey Publishers.

Winterhalder, B. 2001 The behavioural ecology of hunter-gatherers. In P. Rowley-Conwy (ed.) *Hunter-Gatherers. An Interdisciplinary* Perspective, pp. 13-38. Cambridge: Cambridge University Press.

Witter, D. 1992 Region and Resources. Unpublished PhD thesis. Department of Prehistory, Australian National University, Canberra

Zubrow, E. 1990 Modelling and prediction with Geographic Information Systems: A demographic example from prehistoric and historic New York. In K. M. Allen, S. W. Green and E. Zubrow (eds) *Interpreting Space: GIS and* Archaeology, pp. 307-317. London: Taylor and Francis.

FORENSIC ARCHAEOLOGY:

A Textbook

by
John Hunter and Margaret Cox

Routledge

2005

ISBN 0-4152-7311-0
RRP US$115.00
(hardcover)

A RECORD OF STONE:

THE STUDY OF AUSTRALIA'S FLAKED STONE ARTEFACTS

Simon Holdaway and Nicola Stern

Aboriginal Studies Press

2004

ISBN 0-8557-5460-5
RRP A$49.95
(paperback)