Thomas G. Whitley – Gwendolyn Burns

Conditional GIS Surfaces and their Potential for Archaeological Predictive Modelling

Abstract: Conditional GIS surfaces are well known in archaeological applications. Perhaps the most familiar of these are visibility analyses and friction layers (the basis of cost distance evaluations). Archaeological studies however tend to limit the use of such surfaces to purely environmental variables with little explanatory or analytical power. Friction, for example, is usually constructed entirely from a generated slope surface, occasionally with the addition of some physical barriers (such as waterways or perhaps vegetation differences). There are as well, a multitude of potential "conditions" to place upon spatial parameters which include strongly cultural ideas about knowledge, perception, familiarity, territoriality, risk, and other kinds of cognitive behaviours. Here we explore the kinds of "conditions" which could provide a great deal of explanatory or analytical understanding to our GIS archaeological applications, to categorize them into a meaningful framework, and to provide examples from a recent predictive model of how such surfaces can be applied.

Introduction

Not so long ago it was common to see a predictive model which relied on several simple variables; particularly "slope" and "distance to water." Classifications of distance variables especially were quite simply designations such as "less than 100 meters" or "100 to 500 meters." More sophisticated models used more classes, or direct measurement limited only by the resolution of the data (30 meter land units for example). But distance measurements were typically calculated as straight line distances and were performed in a GIS merely by counting the data pixels from a given location (or feature) and translating the result into map distance units.

By using "cost-distance" evaluations, predictive modellers were able to take the first steps toward injecting agency into what would be an impersonal generalization of activity and settlement location choice. A cost distance differs from a distance evaluation only in that it incorporates a moderator of distance; typically in the form of "friction." Today it is much more common to see predictive models (of all varieties) which use cost-distance evaluations rather than straight distance (see van Leusen et al. 2002 for a comprehensive discussion of predictive model methods and approaches). This is the mainstream not just in "academic" predictive modelling, but in heritage management (variably known as CRM, CHM, or AHM) applications as well.

The moderation surface in cost-distance is what we refer to as the "conditional GIS surface". In most applications, friction is usually derived from slope; whether it is a simple transformation of the recorded slope into a relative cost attached to each pixel, or a more complex mathematical calculation of calories expended or projected walking speed. The results are the same but distance is conditional upon the accumulated friction required to reach the objective (water for example). When you consider the nature of conditional variables however, it is easy to see that the "cost" of reaching an objective is moderated by much more than simply slope or the calories expended to cross the land unit. Much more complex issues such as proximity to other resources (i.e. attractors), large scale travel arteries, territoriality, temporal dependence, individual motivations, and completeness of spatial knowledge have a much greater bearing on whether a land unit was ever crossed or occupied for any length of time. Additionally, the nature of the activity has a huge bearing on whether or not we would recognize an occupation (i.e. was any archaeological component ever left behind?). These issues are not easily addressed in a predictive modelling context. By its very nature, a predictive model is meant to be an abstraction either from archaeological site data (in the case of a correlative model) or from hypotheses of settlement choice (in the case of a cognitive model) to a generalization which may be useful for some purpose. But by experimenting with conditional surfaces in a predictive modelling framework we can begin to explore ways in which past people utilized their spatial environments.

Modelling in this fashion has the advantage of building additional levels of agency and cognition into predictive models; making them less generalized and more explanatory. This will also have the tendency to move predictive models away from being the "least common denominator" for all archaeological sites in an area; which we believe undermines their utility (because the least common sites are overlooked, yet may be some of the most significant occupations). With a greater interest in specific conditions there will also be a tendency to build highly specialized models focusing not just on specific time periods but also on kinds of sites (or more properly, kinds of activities). It will also be possible to model activities which leave no archaeological trace or the associations between sites in landscapes of activities (such as resource procurement areas).

Application of more complex modelling methods can seem readily acceptable in an academic context, but may appear to be less so when it comes to heritage management. Heritage management applications of predictive models are driven by the assumption that developers will only pay for models which are inexpensive and make blanket generalizations that can be abstracted to all archaeological occurrences; hence pushing us toward that "least common denominator" approach. However, this is actually a misconception. More sophisticated models are often less expensive because they are not heavily "data-dependent," do not require the "clean-up" of thousands of archaeological site locations, and rely on more robust testing strategies (see Whitley 2004a). There are two independent issues when it comes to cost for predictive modelling: 1) the cost of providing a large and accurate dataset to develop (or test) the model, and 2) the cost of the modelling itself. All expensive predictive models involve high costs of developing the sites dataset, regardless of their theoretical nature. The cost of the modelling itself is often only a fraction of the data development costs.

The primary interests of developers are in getting clearance to begin their construction projects, or in making value choices in selecting possible alternatives. They ultimately have no stake in the theoretical or methodological basis of the archaeological models at all; just their success. Thus, the real driving force behind predictive models is the review agencies who accept their application in a heritage management context. The developers will pay for any model which is successful in getting them their permit, or allowing them to weigh the individual costs of different construction alternatives. Our experience with predictive modelling in US heritage management is that the more sophisticated the understanding of the review agencies, the more likely they are to accept (and require) models that deal with complex phenomena; such as multiple temporal distinctions, different types of activities, and even low (or no) artefact archaeological and historical landscapes. Ultimately, there really is no distinction between academic and heritage management applications of predictive models. All kinds of models are applied in both settings. We need to focus on more explanatory models in both situations. As long as they are successful they will also be accepted.

Example

These same principles briefly described above have been applied to several land-based predictive models that we have conducted recently; one example of which was located in South Carolina *(Fig. 1)*. The following discussion also applies to models we have created in Georgia and Alabama, but only the South Carolina example will be used here. The first consideration is that these models are not correlative in nature; rather they were cognitive models (cf. Whitley 2005). Space limitations prevent a discussion of the details of the models themselves. We prefer here to point out the ways in which we enhanced either the accuracy or the explanatory ability of the model by experimenting with the conditional surfaces.

It has been argued elsewhere that it is incorrect to create models on the basis of assuming that all residents in an area have complete and accurate knowledge of the region, its terrain, and/or resources (see Whitley 2004a, 2004b). We tend to take a global or "god's-eye" frame of reference on an area because we use large scale environmental data that covers huge regions and was typically gathered from aerial imagery or satellite sources. This scale of reference tends to be impersonal and nothing like the real experience of people on the ground (the egocentric frame of reference). This is moderated somewhat by recent applications in landscape archaeol-

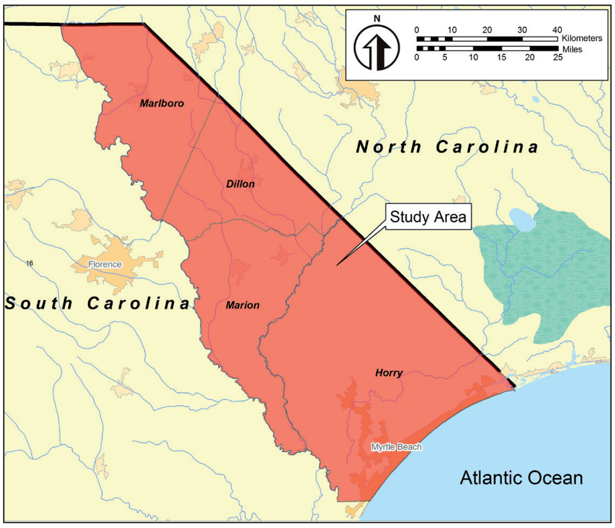

Fig. 1. Study area.

ogy (typically focused on viewpoint or view shed analysis - see Wheatley / Gillings 2002, 200-216 for an overview), which try to recreate that egocentric viewpoint or to develop agent-based approaches (e.g. Gimblett 2002). It can also be suggested that the immersive approaches typical of the virtual reality community in archaeology is struggling to achieve some means of an egocentric perspective on past societies (e.g. Barceló / Forte / Sanders 2000). Cognitive mapping applications in human geography and psychology have also been at the forefront of understanding choice and the egocentric point of view for many years (e.g. Kitchin / Blades 2002). The extensive literature on space syntax (e.g. Peponis / Wineman / Bafna 2001) leaves no doubt that we do have good comparative studies of how people classify their environment and use it on a personal level. However, these approaches are still rarely (if ever) integrated with predictive models, and although we see many interesting applications for them, they are somewhat secondary to this discussion.

As it relates here, we assume that even though we know people do not typically experience their environment from space, that it is a good proxy representation of the accumulated knowledge they have gathered from living in the region during their lifetime and the learned knowledge of their family and peers. Thus, we assume they can make informed decisions about where to conduct their activities regardless of their location in the region. This is also the very (erroneous) basis for testing predictive models against a random distribution of non-site locations (Whitley 2004a). The problem is that activity decisions are not independent operations and they are always made with respect to current locations, or the location of previous activities; because people do not magically appear in the manifold as if dropped from space each time a spatial decision is to be made. Travel is not unlimited and instantaneous, and proximity is the key element to all spatially limited activities. Therefore, by its very nature, spatial knowledge is going to be limited by its proximity to frequent activities; and areas of previous activity are going to be more frequently utilized than other areas farther away. This means the focus should first be put on identifying areas most likely to be used frequently and building a model which incorporates repetitive use areas as the basis upon which to model spatial knowledge.

To do this we developed a series of specific cost-distance evaluations based on what we know (or what is hypothesized) about prehistoric populations in South Carolina. Briefly we defined five primary settlement/subsistence patterns which we believe represent much of the prehistoric and historic period occupations. These patterns are based on real archaeological data from thousands of sites and defined by numerous archaeologists in the region since the early 1900s. These activity-based categories are merely syntheses of the ideas about settlement that have already been devised:

- *Prey-based Nomadism* – This category represents the earliest period of occupation in the region, where nomadic hunters followed migrating groups of large prey, typically in grassland areas, but also in woodlands. It is described as "prey-based" because mobility patterns were largely dictated by the movements of the prey and not tied to spatial boundaries or territories. This period coincides predominantly with the Early Paleoindian period (ca. 14,000 to 9000 years ago).
- *Wide-area Ecosystemic Nomadism* – This category is defined by the transition from large migratory prey to smaller locally abundant resources. Though still nomadic, hunter-gatherers from this time frame exploited large-scale ecosystems for a much greater diversity of species than the earlier time periods. Population pressure was still low however, and there may have been a great deal of regional migration between very different habitats during different times of the year; especially with regard to accessing lithic resources. Temporally, this settlement/subsistence pattern equates with the Late Paleoindian and Early Archaic periods (10,000 to 6000 years ago).

- *Constrained Ecosystemic Nomadism* – This category correlates to a time when population pressure began to create a change in the dynamics of settlement and subsistence in the region. Though still nomadic and still hunting and gathering locally available resources, territories were evolving and becoming constrained. There was likely much greater trade of utilitarian goods, and overwater travel probably became more dominant than in previous periods. In all, this pattern represents the period in which trade and social networks were evolving. The cultural designations associated with this pattern were the Middle and Late Archaic periods (7000 to 4000 years ago).
- *Seasonal Sedentism* – This category represents the period in which previously nomadic people began to re-occupy the same localities year after year in a seasonal round and for longer periods; thus becoming somewhat sedentary. They probably transitioned between larger ecosystems in a fairly regular and predictable pattern. Travelling primarily along waterways (both overland and overwater), they would have maintained close proximity to established trade routes and exploited very predictable local resources in regular and familiar ways. Horticulture became a primary food production method during this time frame; further tying people to the larger more fertile river valleys. The temporal periods associated with this pattern are the Early and Middle Woodland periods (4000 to 2000 years ago).
- *Permanent Sedentism* – This category represents the transition to full scale agriculture and the establishment of permanent villages. Tied very closely to established trade routes, especially within the largest river valleys, political relationships evolved through access to exotic prestige goods. Very complex relationships, territories, and resource exploitation patterns evolved into the Mississippian societies which came to dominate the Southeast. The most significant sites which represent this time frame are the mound centers, located often at the fall line of major river valleys. This settlement/subsistence pattern also includes the later Euro-American occupations through today. Temporally the Late Woodland, the Early and Late Mississippian, and Historic periods (3000 years ago to today) can all be subsumed by this pattern.

Using this framework of settlement/subsistence patterns as a guide, we defined three primary categories of behaviours. Much like the patterns defined above, these should not be assumed to be entirely mutually exclusive, but could easily be shown to have degrees of overlap:

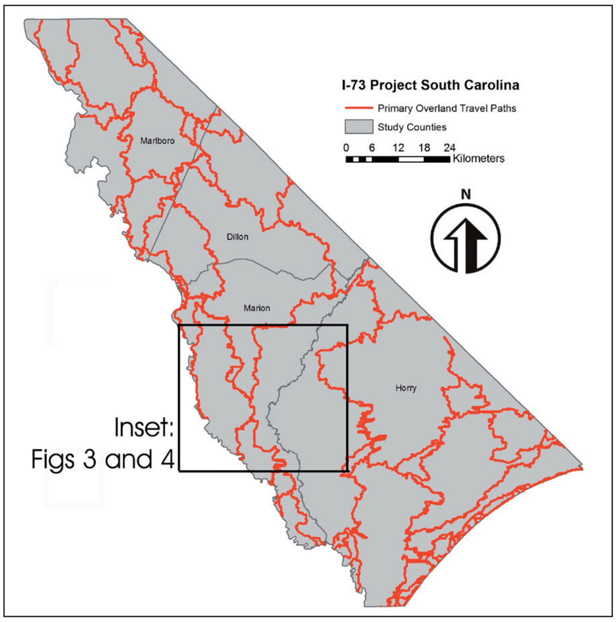

Fig. 2. Overland travel arteries

- *Resource Acquisition* – This includes the hunting or gathering of food species as well as horticultural or farming behaviours, the exploitation of lithic or other resources, and accessing exotic or non-utilitarian items (typically for trade).
- *Domestic/Production* – These behaviours included the establishment of settlements, building dwellings or storage structures, production activities such as lithic tool manufacture, making and firing ceramics, weaving textiles and processing utilitarian resources, etc. They also include cooking food, disposing of refuse, and other domestic activities, as well as manufacturing exotic trade items.
- *Social Interaction* – This includes many kinds of social behaviours such as interaction with neighbours for trade or political purposes, building of mounds or ceremonial centers, territorial boundaries, warfare and other ritual activities.

As an example, we know (or at least we hypothesize) that Prey-based Nomads (in this region) were primarily overland travellers; meaning we can discount over-water travel. They were theoretically tied to the movements of large herbivores. Therefore, large herbivore migratory routes and habitats would be very important spatially limited attractors for settlement (or at least for large herbivore

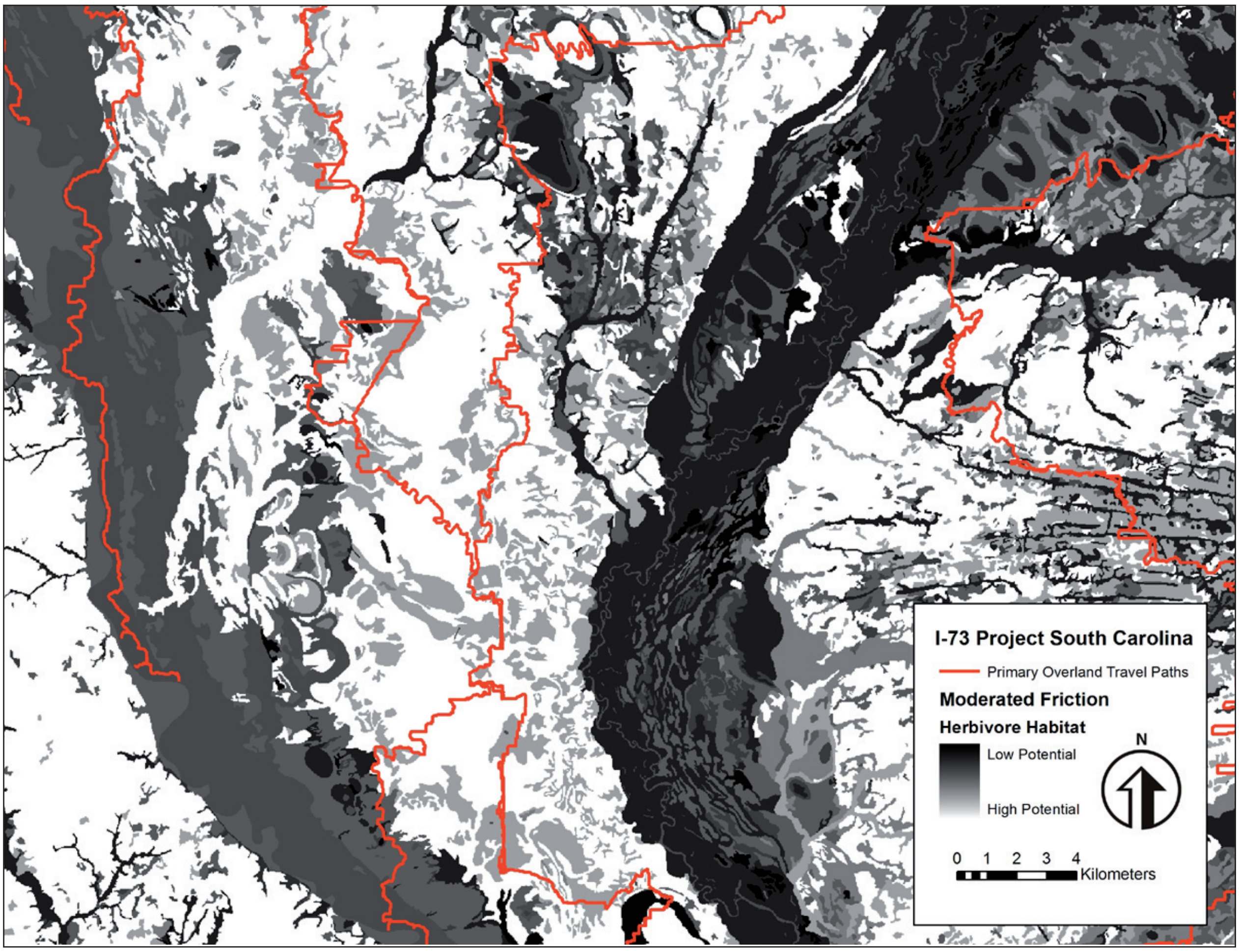

Fig. 3. Detail of GIS conditional surface (moderated for herbivore habitat potential).

procurement activities). We also do not think that they intensively exploited local resources. So the first thing we wanted to do was model large herbivore habitat and migratory routes. Ideally we should have excellent paleoenvironmental records which would allow us to reconstruct exactly what the paleovegetational patterns were at any given time and couple this with an excellent understanding of what kinds of habitat extinct herbivores would have sought out. Of course we do not have that, so we make do with some proxy variables. In this case we know that mammoths and mastodons followed a very large scale migratory pattern; meaning that they crossed great distances in relatively short times. Presumably, Prey-based Nomads would have done likewise and stayed in close proximity to these migratory paths.

To identify the most likely migratory pathways we ran a least cost paths analysis on the entire region using starting and ending points spaced along the edges every 1 km. For the conditional surface we started with a basic relative slope cost but we also added a flow accumulation surface as a proxy representation of water barriers. The flow accumulation model comes from hydrological analysis and is a fair representation of the effort required to cross a river or stream at any given point; the theory being that migratory paths would have self-selected stream fords and avoided crossing large rivers where they were deep and swift. The result was a series of travel arteries which represented the least effort overland paths through the region, and therefore the most likely locations of migratory herbivore travel *(Fig. 2)*.

Hypothetically then, Prey-based Nomads would place themselves conditionally in the environment with respect to their proximity to these pathways. However, their presence would also be conditional on the presence of herbivore habitat in general. Ideally, the best places to find large herbivores would be in the grazing lands placed along these pathways, or at least in fairly close proximity to them. In essence, migratory herbivores would be likely to wander away from their known pathways only

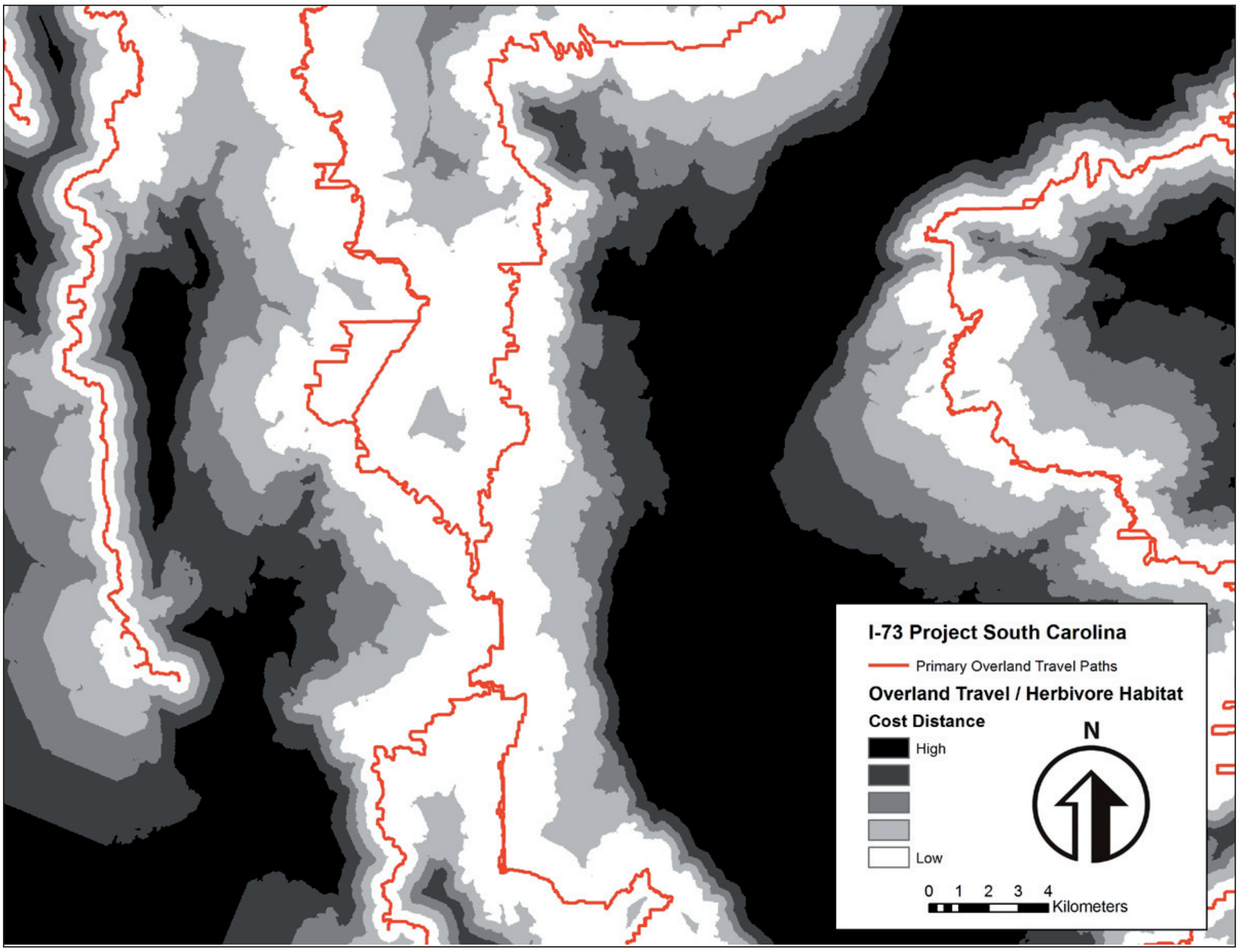

Fig. 4. Detail of a moderated cost-distance evaluation.

if suitable forage were available to them. In other words, the attraction of the pathway is strong where grazing lands are not present and weak where they are, for both herbivores and people pursuing them.

This variation in the attractor is, in effect, another moderator of friction or yet another layer of conditions for the GIS surface *(Fig. 3)*. To model how the pull of the travel artery lessens in the presence of prime grazing habitat, we used soil potentiality values for grasslands, openlands, and light woodlands to give a proxy representation of likely herbivore habitat. In our predictive model then we created a cost-distance evaluation from the primary overland travel arteries that included the same friction costs as before but reduced friction in areas of possible herbivore habitat *(Fig. 4)*. In this way we have theoretically created a proxy representation of the complex ways in which fairly simple hunter-gatherers carried out specific activities. This kind of approach simulates the perspective, the knowledge base, and the motivations of the past people we are studying.

Discussion

We approached the other settlement/subsistence patterns in much the same way; focusing on other conditional surfaces such as over water travel (where we had to distinguish between downstream and upstream friction), pathways for regional resource acquisition, potentialities for horticulture and agriculture, trade networks, exploitation of marsh resources, and the historic land use practices which directed much of the historic settlement. All of these surfaces were combined in different and complex ways in the predictive model, yet still it was a great simplification of the many ways people have used and settled in the region in the past.

Overall we can create very simple predictive models that work reasonably well and serve the purpose for which they were intended without examining closely how people were envisioning and using their environment. Ultimately though these models have no explanatory power. To truly understand the ways in which people were actively engaged in their environment, and the ways in which their frame of reference dictated how activities were spatially distributed, we need to employ more sophisticated means of modelling. This can only be accomplished though experimentation and development of more complex proxy GIS representations.

As referred to earlier, the distinction between academic and heritage management applications seems to drive much of the discussion we have today in respect to predictive modelling (as we witness with the predictive modelling sessions at the CAA every year). The assumption is that few developers will pay for more complex models if they are costlier than simple ones. So how do we move predictive models forward if we are pressured to produce simple, cheap, models? We argue here that in addition to often being less costly in general, complex explanatory models are more meaningful and will produce a higher success rate. We need to differentiate the costs of maintaining a high quality, accurate dataset of sites (which should be encouraged regardless of their use in predictive models) from the cost of modelling itself. In many cases, developers have been paying for data maintenance rather than the actual predictive models. In actual practice we have been achieving high success rates and low costs with explanatory models in numerous settings. This is exactly what is called for in both heritage management and academic applications. Though it may be peripheral to the main point here, simplification does not equate with lower cost, nor with greater applicability. But understanding the application of conditional GIS surfaces to many different situations will be crucial to understanding how people selected site locations in the past, and how we can protect them in the future.

References

Barceló / Forte / Sanders 2000

J. A. Barceló / M. Forte / D.H. Sanders (eds.), Virtual Reality in Archaeology. BAR International Series 843 (Oxford 2000).

Gimblett 2002

R. H. Gimblett, Integrating Geographic Information Systems and Agent-based Modelling Techniques: for Simulating Social and Ecological Processes. Santa Fe Institute: Studies in the Sciences of Complexity (Oxford 2002).

Peponis / Wineman / Bafna 2001

J. Peponis / J. Wineman / S. Bafna (eds.), Proceedings, Space Syntax: 3rd International Symposium, Georgia Institute of Technology (Atlanta 2001).

van Leusen et al. 2002

M. van Leusen / J. Deeben / D. Hallewas / P. Zoetbrood / H. Kamermans / P. Verhagen, Predictive Modelling for Archaeological Heritage Management in the Netherlands. Baseline Report for the NWO (Humanities Section) of the BBO (Stimuleringsprogramma Bodemarchief in Behoud en Ontwikkeling) (Amersfoort 2002).

Wheatley / Gillings 2002

D. Wheatley / M. Gillings, Spatial Technology and Archaeology: The Archaeological Applications of GIS (New York 2002).

Whitley 2004a

T. Whitley, Re-thinking Accuracy and Precision in Predictive Modelling. Paper Prepared for the Computer and Quantitative Applications in Archaeology 2004 Conference, Prato, Italy, 2004.

Whitley 2004b

T. Whitley, Risk, Choice, and Perception: Elements of an Immersive GIS. Paper Prepared for the 69th Annual Meeting of the Society for American Archaeology, Montreal, Quebec, 2004.

Whitley 2005

T. Whitley, A Brief Outline of Causality-Based Cognitive Archaeological Probabilistic Modelling. In: M. van Leusen/H. Kamermans (eds.), Predictive Modelling for Archaeological Heritage Management: A Research Agenda. Nederlandse Archeologische Rapporten 29 (Amersfoort 2005) 125–139.

Thomas Whitley

Brockington and Associates, Inc.
6611 Bay Circle, Suite 220
Norcross, Georgia 30071
United States of America
tomwhitley@brockington.org

Gwendolyn Burns

Brockington and Associates, Inc.
1051 Johnnie Dodds Boulevard, Suite F
Mt. Pleasant, South Carolina 29464
United States of America
innaburns@brockington.org