

Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene

Philip Verhagen · Thomas G. Whitley

Published online: 1 February 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Archaeological predictive modeling has been used successfully for over 20 years as a decision-making tool in cultural resources management. Its appreciation in academic circles however has been mixed because of its perceived theoretical poverty. In this paper, we discuss the issue of integrating current archaeological theoretical approaches and predictive modeling. We suggest a methodology for doing so based on cognitive archaeology, middle range theory, and paleoeconomic modeling. We also discuss the problems associated with testing predictive models.

Keywords Predictive modeling · Cognitive archaeology · Middle range theory · Model testing

Introduction

It has now been more than 20 years since one of the most influential books ever to appear on spatial modeling in archaeology was published: *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling* (Judge and Sebastian 1988). This edited volume of papers centered on the principles and techniques to be applied, and the difficulties encountered, when trying to make a spatial prediction of the potential archaeological record. The wide-ranging contributions in this one volume focused on trying to make archaeological predictions from known observations and how to test them. The

P. Verhagen (✉)

Research Institute for the Heritage and History of the Cultural Landscape and Urban Environment (CLUE), Faculty of Arts, VU University Amsterdam, Amsterdam, The Netherlands
e-mail: jwhp.verhagen@let.vu.nl

T. G. Whitley

Brockington and Associates, Inc., Norcross, GA, USA
e-mail: tomwhitley@brockington.org

ideas outlined still influence the ways in which archaeologists deal with GIS and spatial modeling. They have also, to a large degree, determined the ways in which predictive modeling is currently applied in cultural resource management (CRM).

The volume mainly reflected existing practice in North American archaeology at the time and as such was firmly rooted in the processual tradition. In those days, a quantitative model was used sometimes almost as a goal in itself (see, *e.g.*, Thomas 1978). Furthermore, predictive modeling heavily relied on the ecosystemic approach advocated in New Archaeology (most notably by Binford 1972, 1982). Since then, many researchers working in archaeological applications of GIS and predictive modeling have struggled to come to terms with the ensuing theoretical debate in archaeology between the processual and post-processual schools of thought (see, *e.g.*, Wheatley 1993, 2004; Witcher 1999). Even though some have managed to include aspects of post-processual thought in a quantitative (GIS) framework, many researchers working with GIS and statistics have given up the fight altogether and keep on working in a processual framework—without drawing overt attention to this in their own writings.

As an undesired consequence, the development of predictive modeling has veered away from mainstream archaeological thought and theory and has now become a largely self-contained activity—enjoying reasonable success as a tool for CRM, but not commanding much respect from academic scholars. This has largely resulted from the desire to use predictive models as tools for minimizing field effort rather than for explaining the differential spatial patterning of archaeological sites. Although the debate is far from conclusive regarding the benefits of predictive modeling in the world of heritage management, it is clear that many current applications in CRM are often simplistic and intended by non-archaeologist land managers to be cost-saving rather than explanatory.

Over the past 10 years, a few practitioners of predictive modeling have been trying to work on alternatives to current practice (see van Leusen and Kamermans 2005; Kamermans *et al.* 2009). Up to now, their efforts have not led to a real breakthrough in thinking about predictive modeling outside the small circle of scholars involved. In this paper, we want to lay out the background to the issue, restate the case for integration of archaeological theory and predictive modeling, and develop a methodology for doing so. While there are obvious cost-saving benefits for even simple predictive models in the area of selecting alternatives or even motivating land managers to fund protective measures or place an emphasis on maximizing interpretive results, we specifically want to focus on the process of theory-based modeling rather than on the resulting models and the uses they are put to. We hope that this will point the way out of a debate which we feel has been unduly polarized along the lines of CRM *versus* academic research, as we are convinced that predictive modeling can be a useful instrument for both fields of application.

A Short History of Predictive Modeling

The roots of predictive modeling can be largely traced back to the rise of processual or New Archaeology in the late 1960s, even though archaeologists have always been interested in issues of where sites are located and why. The pioneering work of Willey (1953) in the Viru Valley, as the genesis of directed settlement studies, greatly

influenced the development of later analyses and both the theory and methods chosen to build predictive models. But the realization that the location of human settlement is closely related to characteristics of the natural environment, coupled with the processual emphasis on quantitative methods, was what enabled the formulation of statistical models for the prediction of site density in areas where no archaeological sites had yet been found.

While the term “predictive model” itself can be traced back to publications of the early 1970s, it is only in the second half of that decade that predictive models started to be produced on a larger scale in the USA. Interestingly, the general methodology of predictive modeling was developed long before the first affordable GIS software became available in the mid 1980s. Seminal research on issues of predicting site location was carried out by the Southwest Anthropological Research Group in the early 1970s (Plog and Hill 1971). It also seems that the first use of multivariate statistics to predict site location was already applied by 1973 in the former British Honduras (Green 1973).

By the 1980s, two primary lines of research were being called predictive models: first, a largely theoretical series of models based on the use of ecosystemic structures and relationships to identify spatial suitability (*e.g.*, Jochim 1976, 1981; Bettinger 1980) but which had no substantial quantitative evaluation of a spatial area (in the same sense as we would expect in a GIS) and, second, a series of models that were based on extrapolating the environmental variables from the landscape in a quantitative fashion and building correlative statistical summaries that could be applied in unsurveyed areas (*e.g.*, Kvamme 1983, 1984, 1985; Parker 1985) yet which were by no means easy to produce. The earliest models produced by Kvamme (1983, 1984), for example, were constructed using an advanced pocket calculator that could be programmed to perform the necessary calculations. It is therefore inaccurate to assume that GIS has been the main factor influencing the initial development of predictive modeling, although it certainly helped its further proliferation in the 1990s.

In later publications, the statistical extrapolation methods developed in the early 1980s are often referred to as “inductive” (Kamerians and Wansleben 1999), although the terms “correlative” (Sebastian and Judge 1988) and “data driven” (Wheatley and Gillings 2002) are also used (Fig. 1). This form of modeling compares known site data, usually within a controlled survey area, with “environmental” datasets like distance to water, soil type, and slope, and then extrapolates the correlations found to areas where no site information is available, usually by means of logistic regression (see, *e.g.*, Warren 1990). Archaeological theory on site location preference only plays a very limited role in these models. The variables analyzed and used for prediction are obviously thought to have some relation to site location preference, but little or no attention is paid to ideas on how people may have used and perceived the landscape in the past.

This way of modeling is contrasted to a “deductive”, “explanatory,” or “theory driven” approach, in which archaeological site information is not used to look for correlations but only for testing purposes (Fig. 2). These models are based on hypotheses of settlement location preferences and are modeled using relatively simple GIS techniques like weighted overlays. Dalla Bona (1994), for example, used theories about the subsistence practices of Native American hunter-gatherers in Ontario to model (among other things) potential moose habitats as attractive zones

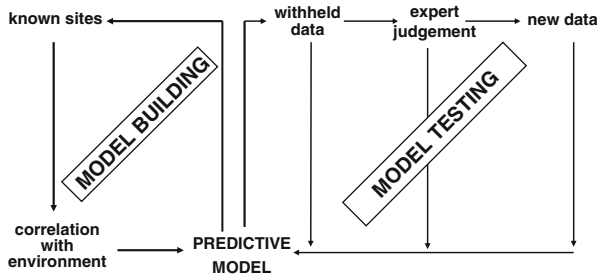


Fig. 1 The procedure used for inductive predictive modeling. A statistical comparison of archaeological data and “environmental” variables is used to create a predictive model. This model is then tested either through statistical methods using withheld or new data or by means of peer review (expert judgement). After Kamermans and Wansleeben (1999)

for site location. Initially, these explanatory models were mainly used in regions with a relative absence of site information. Later they were also applied in areas where the available site information could not be reliably related to survey activities (see, *e.g.*, Deeben *et al.* 2002). In many cases, these models are relatively unsophisticated, to the point of being called “intuitive” or “expert judgment” models. While these may successfully be combined with statistical methods (Verhagen 2006; van Leusen *et al.* 2009), current practice still relies on either inductive or deductive modeling, not on approaches where both are combined. We want to stress however that the dichotomy between the “inductive” and “deductive” arose in the late 1970s as a historical development and not necessarily as two methodological schools of thought, and as such they should not be thought of as mutually exclusive frameworks.

At this point, it may be useful to clarify our own position regarding the definition of predictive modeling in archaeology. A basic definition was already introduced by Kohler and Parker (1986, p. 400):

Predictive modeling is a technique that, at a minimum, tries to predict “the location of archaeological sites or materials in a region, based either on a sample of that region or on fundamental notions concerning human behaviour”

Following this definition, we will only speak about *predictive* models if they result in a quantitative estimate of the probability of encountering archaeological

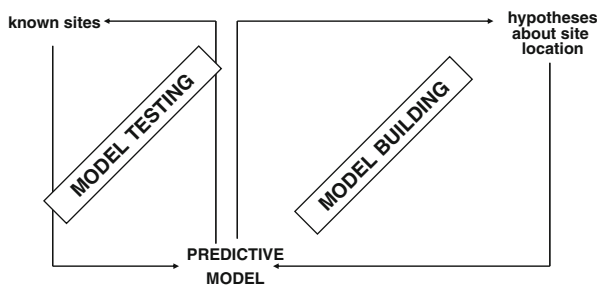


Fig. 2 The procedure used for deductive predictive modeling. The model is created on the basis of hypotheses on site location preferences. In many cases, these are no more than “educated guesses.” The model can easily be tested with statistical methods as the archaeological data set is not used for building the model. After Kamermans and Wansleeben (1999)

remains outside the zones where they have already been discovered in the past (Verhagen 2007a, p. 14). This is regardless of whether the model is used for CRM purposes or serves as a testable hypothesis for further scientific research. Prediction is an essential outcome; “models” without predictions are better categorized under the header of site location analysis. But since we are confronted, in the process of setting up the models, with numerous theoretical, technical, and methodological issues of spatial analysis in archaeology, it may not come as a surprise that Altschul *et al.* (2004) stated that many archaeologists now believe that spatial analysis (or “using GIS”) equates to predictive modeling.

The principle of statistical extrapolation, particularly as it was used in correlative models, proved to be useful for CRM purposes in the USA. After the introduction of the National Historic Preservation Act in 1966, federal agencies, confronted with the question on how to deal with their responsibility to “identify historic properties on their lands (...) and to record such properties when they must be destroyed” (King 1984, p. 116), generated a demand for what was initially called “predictive survey” to limit the costs associated with identifying and managing such resources. Federal agencies with these management responsibilities were able to fund such studies and helped initiate their proliferation. But while the CRM industry provided the monetary backing for predictive models and the goals expected in a CRM context were driving the format of their development and application, the theoretical debate and the discussion regarding acceptance or rejection of predictive models as a whole was still being framed within the scientific community.

This is perhaps best seen in the influence of landmark syntheses and theoretical articles that look at spatial decision-making. Aside from the hunter–gatherer studies of Jochim (1976, 1981) and Bettinger (1980) already alluded to, an excellent example is the work of Limp and Carr (1985), which initiates decision-making as an interpretive topic in archaeology and provides a forerunner for both agent-based and GIS modeling. Kohler and Parker (1986) laid the groundwork for a definitive look at the practice of predictive modeling itself, while Judge and Sebastian (1988) followed it up and fleshed it out. Other landmark studies that definitively affected the course of predictive and other forms of GIS modeling include Allen *et al.* (1990—particularly the individual chapters by Warren, Altschul, and Savage), Renfrew and Zubrow (1994), Harris and Lock (1995), Lock and Stančić (1995), Aldenderfer and Maschner (1996), and Wescott and Brandon (2000).

The use of predictive models in CRM has engendered both enthusiasm and criticism. In the USA, Canada, The Netherlands, and to a lesser extent in Germany, the Czech Republic, and Australia, predictive models are routinely used to obtain an assessment of site density in an area that might either be facing a direct threat of disturbance or in a planning zone where policies on CRM need to be determined. In other countries, predictive models are at best seen as interesting sources of information when preparing for survey, but in some cases predictive modeling is outrightly rejected as a tool for CRM. Resistance has been particularly strong in the UK and France. The argument is often voiced that the models are by definition unable to predict the location of *all* archaeological sites (see also Altschul *et al.* 2004). Especially in areas that are predicted to be of “low archaeological potential,” predictive models are in practice used to recommend a less intensive survey strategy, and in this way sites may be overlooked and lost forever.

While it is illusory to think that we will be able to detect all archaeological remains in an area without stripping it completely, the practice of targeting areas for survey on the basis of predictive models is still regarded by some as almost sacrilegious (see, *e.g.*, Wheatley 2004). However, practice shows that when predictive models are not used, there will be less opportunity for archaeologists to influence spatial planning in the early stages. A predictive map can be a very powerful and useful instrument for the protection of the archaeological heritage—when used wisely. An alternative analysis approach can, for example, be employed for certain types of CRM projects. It has been shown to lead to significant financial savings yet is not as controversial because it does not recommend differential survey strategies. In this application, a predictive model is still used to determine which areas are more likely to produce archaeological sites and which are less likely to do so, but the field effort is not changed between high- and low-potential zones. Instead, the same level of effort is employed throughout, but where there are several different options the planners are forewarned about the specific costs which are likely to be entailed with each alternative. Therefore, they can choose the one which will probably be less costly in terms of archaeological survey and later testing and mitigation costs can be anticipated. This approach is commonly used for large-scale highway planning purposes, and especially in the case of well-established and well-defined models (such as the Minnesota Model; Hudak *et al.* 2002) it has been used with financial success and is supported by the review agencies involved.

The conflict between the theoretical goals of scientific research and the practical matters of budget management by the efficient targeting of survey resources is almost pre-ordained by the focus of both research and CRM on the archaeological “site” as the base unit of analysis. The site is seen as the most significant locale for both research and protection because of its accumulation of material debris. A site’s level of protection, however, is typically based on its potential for future research (whether it was located with a predictive model or not). A predictive model which recommends a less intensive survey strategy for some areas is then merely classifying some sites as being of lesser importance than others (*i.e.*, those that are predicted by the model are presumed more important for preservation and hence future research). But this is little different from the classification strategy used by all archaeologists when they choose to ignore certain kinds of site (such as non-diagnostic lithic scatters) or even “non-site” locations in general. It is also the strategy used when they develop intuitive models for targeting research survey projects.

The entire notion of a “site” presupposes that there are parts of the landscape that are less important for research or preservation at all—presumably those areas with no accumulation of artifactual debris. But is it correct to assume that prehistoric people would have classified their landscape the same way? Is it not logical to assume that vast areas of the landscape were important to prehistoric people for the resources they provided (among other things)? Yet these areas could never be classified as archaeological sites or undergo protection from development or eradication merely because no artifacts accumulated, the material remains were all perishable, or these are otherwise not detectable today. If we focused on understanding how people inhabit and use their environment from a cognitive perspective and not by where they accumulate their detritus, then the “site” becomes

merely an area where a certain kind of behavior led to the accumulation of debris or features—not the entirety of research potential. This becomes particularly important as predictive modeling matures theoretically.

Predictive Modeling as a Scientific Endeavor

Understandably, given the traditional theoretical training of today's archaeologists and the powerful funding mechanisms of CRM, the focus of research in predictive modeling has been on its utility for cost savings and the improvement of existing techniques and data sets rather than pushing the envelope, so to speak, of theoretical research. This does not imply an absence of innovation however, especially in the realm of statistical analyses. The suite of applicable statistical techniques has, amongst others, been extended to Bayesian statistics (Millard 2005; Finke *et al.* 2008; van Leusen *et al.* 2009), Dempster–Shafer modeling (Ejstrud 2003, 2005; van Leusen *et al.* 2009), and resampling (Verhagen 2007b). The importance of bias identification and reduction in both environmental and archaeological data sets has long been recognized (Ducke and Münch 2005; Verhagen 2007b), as well as the relevance of using geo-archaeological data for the reconstruction of past landscapes (*e.g.*, Zeidler 2001; Peeters 2007). However, relatively little attention has gone to the positioning of predictive modeling within a scientific research framework (but see Whitley 2004, 2005).

An extreme position in this might be taken by some practitioners of CRM who feel that the work done in CRM is emphatically *not* science, although survey and excavation should be done according to scientifically proven and approved methods. In this way of thinking, it does not really matter what kind of methods we apply as long as we can be certain that they produce a sufficiently reliable result. At the other end of the spectrum, we find academic researchers who feel that science (in its meaning of “hard”, quantifiable science) is alien to the core business of archaeology: the interpretation and explanation of past societies. From this point of view, traditional (either inductive or deductive) predictive modeling provides an easy target for criticism as it provides no or insufficient understanding of how and, most importantly, why settlement patterns have come about.

Dobres and Robb (2005) describe the “normal” archaeological research process as starting with the formulation of theory, which then leads to the choice for a method(ology) to collect data. In this scheme, models are potential end products of research together with explanation and interpretation. While they object to this linear approach and advocate what they call an interdigitating research practice, it is clear that they consider this to be the standard way for archaeologists to do research. It also explains why the predictive models are regarded with so much suspicion. If we expect models to be on equal footing with interpretative and explanatory accounts, then we will be almost inevitably disappointed in them, and using them as starting points (like we do with predictive models), rather than end products, seems inherently incorrect.

Modeling (of any type) can be better considered within a slightly different generalized scheme of scientific research (Fig. 3, somewhat adapted after O’Sullivan and Gahegan 2007) that starts with the collection of data (acknowledging of course

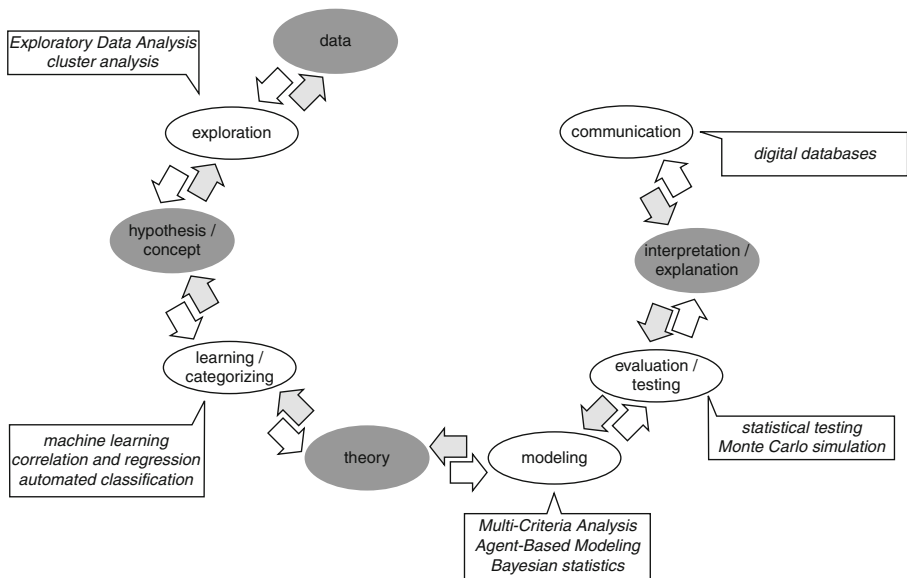


Fig. 3 The position of modeling in the scientific research process. Quantitative methods and models can play a role in almost any stage of scientific research but are best applied in the stages where theory is developed and made suitable for testing. The separation between hypothesis and theory however may not always be as clear as suggested. Adapted after O’Sullivan and Gahegan (2007)

that data are never collected without a research question in mind). The first activity is then a phase of exploratory data analysis: what are the features of the data set that could possibly answer our research questions? In this phase, several exploratory statistical techniques can be useful to detect patterns and perform classifications. On the basis of the patterns that suggest themselves in this exploratory phase, a second step is then the establishment of a “true” theory, *i.e.*, one that transparently specifies the (quantitative) relationships between the variables that are thought to be responsible for the patterns detected. Inductive statistical models can then be used to generalize this theory to unknown instances, but generalizations obviously can be made on the basis of logical arguments as well—and this is what we usually understand by deductive modeling. The results of these generalizations should then be open to scrutiny; a testing phase is needed in which new data are collected to see whether the theory holds in the light of new evidence. If this proves to be successful, then theory, model, and evidence combined can be published and represented with maps and graphs; this is the basis for rhetoric, the core of scientific debate and publication. In this scheme, modeling is only a tool for the generalization of theoretical concepts. In this way, it also constitutes a method for opening up a theory to testing and as such does not preclude backtracking to the original theory and data when the model fails to deliver.

It will be clear that the exploratory (theory building) and testing phases in this scheme are the most problematic for predictive modeling. In many instances, the data used for predictive modeling were never collected with the explicit aim of constructing a predictive model from it, creating all kinds of bias that are undesired from a statistical (testing) point of view. Archaeological survey data are the most

notorious in this respect, but the same holds true for many of the environmental “variables” that are customarily used for predictive modeling, like digital elevation models and soil maps. First of all, the accuracy and amount of detail of these can vary greatly; digital elevation models are especially notorious in this respect, and any errors in the original elevation maps will also have repercussions on the accuracy of derived data like slope maps, cost surfaces, and viewsheds (see, *e.g.*, Nackaerts *et al.* 1999). Furthermore, the classification schemes used for soil maps are intended for modern agricultural land suitability valuation purposes. While the modern view of land suitability may bear some relation to the (pre-)historic perception of land quality, this is by no means a straightforward relationship (see, *e.g.*, Favory and van der Leeuw 1998, pp. 278–284). Thirdly, landscapes are dynamic systems that are subject to constant change. Cartographic representations of the current elevation, soil type, and other variables will therefore not necessarily reflect the situation in the past (see, *e.g.*, Zeidler 2001).

Furthermore, the theory building that should go on between the phases of data exploration and extrapolation is currently often neglected. Theory, when it is considered at all, is either assumed to be existing and only needing extrapolation or it is reduced to specifying rather simplistic (ecological) cause–effect relationships that can be fit to the available environmental data sets. For example, the assumption that prehistoric farmers would have preferred fertile land for agriculture is hardly contentious. However, when asked where this land might have been, how much of it would have been necessary to support the population, and where settlements might have been located with regard to the accessibility of this land or any other resources, theory readily gives way to the basic practical notion that certain soil types are more productive than others and therefore the corresponding units depicted on the soil map will have supported more settlements. This assumption is made without regard to whether the prehistoric farmers had the ability to measure and track slight differences in soil fertility across the landscape and place settlements accordingly.

However, translating theory into a (predictive) quantitative model is not always easy. Tschauner (1996) states that most archaeological theory building can be characterized by generalizations in two forms: “general laws” and “statistical laws.” The first type of law can be recognized in statements such as “ritual is ‘important’ to leadership, and institutionalized ritual is an ‘effective means’ of protecting and legitimating power” (Tschauner 1996, p. 11). These are qualifying statements that are usually arrived at through logical deduction. The second type of law usually can be recognized by reference to frequencies and numbers, for example, when it is said that ceremonial rituals are “frequently associated” with the establishment and reiteration of relationships between people, objects and images. These “statistical laws” however should not be confused with true statistical statements about frequencies and associations as they might easily be based on general impressions rather than robust statistical analyses.

Note that neither “general” nor “statistical” laws necessarily imply the specification of causal and/or quantitative relationships. Hence, no probabilities can be attached to these theories other than the ones suggested by the archaeological data set themselves. And while predictive models based on such theories may then seem to produce reasonable results from a CRM perspective (like predicting 70% of the archaeological site sample “correctly” or rather placing them in a zone of high

archaeological potential), there is no mechanism available to test the *theoretical* assumptions of the model. The absence of causality in such statements would draw considerable criticism from philosophers of science (e.g., Salmon 1971, 1998).

So, while predictive modeling is from a scientific point of view a perfectly laudable and worthwhile enterprise that can be used to extend theories of spatial patterning into testable generalizations, in practice it stumbles when dealing with developing and testing site location theories in archaeology. The absence of well-developed causal explanations tends to relegate archaeological predictive modeling to either a vague intuitive impression of “high sensitivity” areas or a series of “so what?” correlative statements. But how do other disciplines then deal with this?

Predictive Modeling Outside Archaeology

In fact, other disciplines do not fare much better, although some of these might seem to be better equipped with tested and testable theories. We can make a basic distinction between disciplines that try to predict on the basis of mechanistic and/or biological processes, like geology and ecology, and those that are dealing with issues of human decision-making, like economics and politics. Archaeology in that sense takes an intermediate position as it combines elements of both.

Geology is a prime example of a discipline where human decision-making issues are completely absent. Predictive modeling is mainly used as a tool for mineral exploration purposes. The available literature is relatively limited and is mainly concerned with the use of weight-of-evidence techniques (a spin-off of Bayesian statistics; Bonham-Carter 1994; Raines 1999) for predicting mineral occurrences at the earth's surface. More recent papers also discuss the use of fuzzy logic and logistic regression in these models. Geological research in general however tends to focus more on developing complex 3D-models for mineral resources prediction at large depths. For these applications, current 3D GIS is not the best tool available due largely to the problems of dealing with volumetric data in three-dimensional space. The weight-of-evidence models are data-driven, and concerns about the effects of exploration bias on the modeling results can be found in some publications (e.g., Coolbaugh *et al.* 2007).

Quite remarkably, theoretical concerns about the formation processes of mineral ores seem to be completely absent: the motivation for choosing the independent variables that predict ore occurrence is flimsy to say the least. It is therefore hard to judge whether this reflects a true lack of understanding of the processes that influence mineral distribution or whether these variables are considered to be almost self-evident predictors. Oreskes (1998) however clearly indicates that the geosciences experience similar problems with regard to model building and testing as the social sciences in the absence of sufficiently accessible and measurable data. She even states that “geological, biological and ecological models have no historic track record of predictive success at all” (Oreskes 2003, p. 24).

Ecology may in fact be the discipline where the closest parallels to archaeological predictive modeling are found. Ecologists depend on fragmented and uncertain information on the spatial occurrence and abundance of animal and plant species, and the factors determining species distribution in different ecosystems are only

partly understood. Predictive modeling in ecology is also closely connected to management issues related to wildlife conservation and protection programs and is therefore strongly driven by a need for accurate predictions.

Since ecologists are dealing with the physiological responses of animal and plant species to environmental conditions, one might expect them to be in a better position to use quantitative deductive modeling procedures than archaeologists are. Nevertheless, much of the theory underpinning ecological predictive models seems to be highly generalized, and still inductive modeling using linear or logistic regression methods prevails (for some examples, see Hoving *et al.* 2004; Mathys *et al.* 2006; Zimmermann and Breitenmoser 2007; Hengl *et al.* 2009). In recent years, random forest models—a machine learning technique—have become more popular (*e.g.*, Cutler *et al.* 2007; Peters *et al.* 2007). Huston (2002) states that the theory of community ecology (which he sees as the body of ecological theory most relevant to species occurrence modeling) “lacks a rigorously tested and widely accepted theoretical framework.” As it is at the same time described as a “broad subdiscipline spanning population and ecosystem ecology and including a range of processes operating from the molecular to the continental,” this is perhaps not very surprising. Ecology is also dealing with highly dynamic processes in which species compete for space and in which seasonal rhythms play an important role. These processes are usually approached through some form of quantitative deductive modeling, without much empirical data to support the theoretical assumptions used. Agent-based modeling, while frequently used for simulating and understanding ecological and evolutionary processes, is not applied extensively for predictive purposes.

Most reported studies suggest that reasonable or even high success rates are achieved in predicting the habitats of the species considered. The figures cited however are not usually in excess of the predictive success of many archaeological predictive models.

Ironically, predictive modeling is also widely used in another form of CRM—customer relationship management (*e.g.*, Rygielski *et al.* 2002; Verhoef 2003). It is a form of marketing (or market analysis) which helps companies target consumers based on predictions derived typically from personal data, demographics, and past purchases. Such economics-based predictive models (also called predictive analytics in this field) analyze the likelihood of both sales and customer retention. One example is healthcare insurance providers who employ predictive modeling methods to proactively avoid catastrophic costs associated with high-risk group members and health crises. This may include projections based on genetic correlates with deadly (or costly) diseases. Models of this kind are rarely theoretically described, only occasionally have spatial characteristics, and are based on many of the same principles and assumptions defined above. They include sophisticated methods and even have a dedicated software (*e.g.*, DTREG: <http://www.dtreg.com>). One peculiarity not generally employed in archaeology includes “uplift modeling,” which is a type of analysis based on the change in probability resulting from a recorded action. Such models are sequential or iterative, and causality is expressed much more explicitly than a vaguely correlative statement like: “a diet rich in fiber may help reduce the risks of some kinds of cancer.” Uplift models are typically utilized in analyzing rates of customer retention after changes in product design, consumer economic conditions, or offered services. In general, intuitive or

correlative predictive models are widespread in the healthcare, actuarial, and economics disciplines and are fairly well accepted without much detailed discussion on the validity of their statistical or theoretical assumptions.

So, despite the wide variety of disciplines that apply predictive modeling for many kinds of analysis (of which we have not attempted to provide a comprehensive overview), it seems that the theoretical underpinning of predictive models is equally challenging whether we are dealing with mechanistic and/or biologic processes or with human decision-making. What makes archaeological predictive modeling perhaps even more demanding is the fact that it is a discipline in which both lines of inquiry can play a role. Archaeological predictive modelers would therefore do well to keep an eye open for alternate methods applied in other sciences.

The Role of Post-processual Theory

The post-processual school, initially developed in the early 1980s in response to the seemingly anti-humanistic, ecosystemic, and deductive–nomological philosophy promoted by the New Archaeology, has dominated much of the archaeological theoretical discourse over the past 25 years, especially in the UK. Even though the heated debate between the two schools now seems to have become somewhat outdated (see, *e.g.*, Hegmon 2003; Johnson 2006) and not many archaeologists nowadays would profess to be adherents of the “hard-core” variants of either the processual or post-processual paradigm, the effects of this debate (both good and bad) are still felt.

Three important characteristics of post-processual archaeology are worth emphasizing from the perspective of predictive modeling. Firstly, post-processual theory usually strongly downplays the determining influence on human behavior of both the natural environment and cultural “systems” of social organization. Instead, the role of individuals (“agents”) is considered to be the driving force of all archaeological patterning and as such should be the prime focus of archaeological research and interpretation. This aspect of post-processual thinking is taken aboard by many archaeologists nowadays, even though not all would see it as their principal research goal. Especially North American prehistoric archaeologists are still reluctant to see agency as a useful concept for their research. Furthermore, post-processualism emphasizes the importance of individual experience. This has led some archaeologists to fully embrace phenomenology and try to engage the mindset of prehistoric people by simulating an immersion in their environment, by installing doorframes in the landscape, for example, in order to understand what it would be like to look at the landscape from inside a dwelling (see, *e.g.*, Tilley *et al.* 2000). These approaches are much more controversial and do not seem directly relevant to predictive modeling, although they do suggest that archaeologists are beginning to understand the need to recreate past cognitive processes.

Secondly, post-modern philosophy maintains that all scientific knowledge is a social construct. Archaeological knowledge is continually reconstructed by archaeologists and therefore fluid, relative, and subjective (Shanks and Tilley 1987). This means that the creation of multiple interpretations of the past is an integral and indispensable aspect of doing archaeology. However, it is also taken to mean by

some that looking for objectivity and scientific verification is pointless (Hodder 1991; see also Fleming 2006 for a critique), which seems paradoxical as these ways of constructing knowledge should then be at least on a par with other interpretations. This is probably the aspect of post-processual philosophy that has inspired the most criticism as it is easy to show that this anti-scientific position mistakes using scientific methods for scientism and positivism—which was indeed a characteristic of the New Archaeology (Bell 1994). The not-so-hidden agenda of early, radical post-processualism was therefore not to promote multiple interpretations in which “objectivist” and “distanced” views can play a role next to more “engaged” ways of looking at the past but an outright rejection of the empiricist and “scientistic” views.

A third conspicuous characteristic of post-processual archaeology, directly following from this anti-scientific stance, is the importance it attributes to narrative in archaeological interpretation. Tilley (2004, p. 225) even went so far as to describe phenomenological writing as “a metaphorical work of ‘art’ for which we make no apology.” Obviously, this means that maps, graphs, and tables are to be avoided in these writings as these are thought to be attempts at objective, “cartesian” descriptions—even though anyone who has ever tried to make maps knows that these constitute specific interpretations. And, in fact, quantitative models are representations as well and have even been described as “a work of fiction” (Cartwright 1983, p. 153). In this sense, all archaeological interpretations are “works of art” whether they contain specific graphical representations or not and whether they evolve from processual or post-processual theorists. However, some might amend that argument by stating that the “art” is in convincing your colleagues that your interpretations are closer to the original thought and meaning of the people in the past than other interpretations. The scientific method is merely the process by which support can be generated amongst your colleagues and maintained for your specific interpretations. The post-processual perspective would seem to reject the notion that anyone other than the interpreter needs convincing and that all interpretations are inherently equally valid, with or without collegial support.

Post-processualism has not been without its critics; yet, it has certainly led to a decrease in the overall interest in quantitative approaches to archaeological questions, and those interested in the contribution that “hard” science may have to offer to archaeology have often found it difficult to wage a successful defense (see, e.g., Pollard 2004). So, the research framework that was sketched above, in which statistical methods and quantitative modeling form an integral part of “doing science,” is not the way in which many archaeologists nowadays would see their discipline. Instead, scientific methods and results are usually seen as auxiliary to archaeological interpretation, like providing better chronological resolution in the case of 14C-dating or aiding survey in the case of geophysical measurements. GIS is routinely applied as a tool for “science” in archaeology, but it is almost always used in either a straightforward data mapping overlay (often for data mining) or as a method to achieve a low-level analysis, such as artifact densities, simple pattern recognition, or regional settlement summaries. Quantitative models are not necessarily regarded as useful heuristic tools to develop, improve, and test an archaeological theory.

This is not to say that no one has tried to look at it that way. In fact, one of the reasons why GIS attracted much interest in the late 1990s was that, in spite of its

“processual” roots, it promised to be a tool for producing a form of quantitative phenomenology through the use of cost-distances and viewshed calculations (see, e.g., Witcher 1999; Llobera 1996, 2000, 2001, 2003). It has also been used to create spatial models of land use from an agency perspective (e.g., Robb and van Hove 2003; Trifković 2005) through the concept of taskscape (Ingold 1993, 2000).

Most attempts to combine processual and post-processual thought in modeling however cannot be considered very successful. Cognitive archaeology, for example, was mounted by Renfrew and Zubrow (1994) to counter the early post-processual critique on New Archaeology by explicitly including social (cognitive) aspects into a systemic view of human behavior. However, it was not embraced by many and seems to have faded into oblivion. The basic precepts of cognitive archaeology though were once again indicative of the understanding that we may never grasp the entirety of what people in the past thought or why they made certain decisions but, through the process of logical deduction, we could build models for recreating some aspects of their cognitive landscape or culture. The processualists opposed this with the apocryphal statement that the only time an archaeologist enters the mind of their subject is when the trowel slips while excavating a burial. The post-processualists also opposed the same notion by rejecting the need for logical deduction—there was no need to convince, thus no need for logic. Unfortunately, cognitive archaeology arose at a time in which the two extremes of archaeological theory were at their most divisive, and no real conciliation between them was sought or expected by either side.

Complexity theory has been suggested as a way out as well (van der Leeuw and McGlade 1997; Bentley and Maschner 2003). Recent developments in agent-based and non-linear modeling show that this branch of theory is far from dead (Kohler and Gumerman 2000; Beekman and Baden 2005; Kohler *et al.* 2007; Dean *et al.* 2006; Axtell *et al.* 2006; and Gumerman *et al.* 2006). In fact, it incorporates two important characteristics of post-processual thought in a quantitative, dynamical modeling framework by explicitly including the role of agents in creating (spatial) patterns and by allowing for multiple model outcomes and thus multiple interpretations. Modeling in this framework is primarily used as a heuristic tool. It is however not usually seen in a post-processual light (but see Hegmon 2003) and has failed to make a breakthrough into mainstream archaeology up to now not only partly because of the technical difficulties involved but also because of the simplicity of the rules employed. Critics of complexity theory assume that complex human agents cannot be modeled with relatively simple parameters and decision-rules.

We can distinguish two varieties of agent-based models: those which incorporate agency as a theoretical concept in how the model operates and where individual or group decision-making is integral to its explanatory power (e.g., Wilkinson *et al.* 2007; Cleuziou 2007) and those which employ programmed cellular automata to run through an iterative process on a spatial manifold according to general rules of behavior, typically in a simulated environment (e.g., Epstein 2006) but sometimes in a GIS (*cf.* Gimblett 2002). These can be seen as “passive” or “active” agent-based models, respectively.

If we want to theorize about the cognitive landscape or the variations in how past people made decisions, agency should be incorporated in a predictive model at least on the theoretical level. However, if we can simulate the cognitive elements involved

in spatial decision-making, then cellular automata can be programmed to evaluate the relevant criteria according to behavioral rules. This is exactly the approach taken by researchers in the American Southwest (*e.g.*, Kohler *et al.* 2007; Dean *et al.* 2006; Axtell *et al.* 2006; Gumerman *et al.* 2006) and is the basis for the models of Mesopotamian urbanization used by Wilkinson *et al.* (2007).

Because agent-based modeling is interested in examining the spatial basis of decision-making, it should be used to predict both intentional and unintentional actions and the ways in which those actions come about. An agent-based model needs to rely not only on intentional actions (programmed rules for cellular automata) that fit with the spatial conditions at the time (the cell values at each iteration of the model), but they must also illustrate each component of the process, realistically resolve conflicting information, and generate outcomes from the limited and conditional knowledge that an agent would have had.

This implies that an explanation of spatial behavior should incorporate a “perspective” or a representation of the perception held by the agent. An understanding of the causal processes entailed by the model does not come from a global view but one conditioned by the costs and benefits of an action, along with the knowledge, confidence, and risks involved in taking that action. This is perhaps more complex than cellular automata can currently be programmed for in a simple spatial manifold, but it is not out of the realm of the potential, particularly when applied to real landscapes in a GIS.

From the foregoing, we conclude that the main problem with (hard-core) post-processualism with regard to modeling is found in its rejection of scientific methods, in the mistaken assumption that using these will automatically lead to “scientism” or “positivism”. We embrace the post-processual emphasis on pluralism, as the only way in which archaeology as a science can move forward is by looking at problems from different angles (see also Wylie 2002, pp. 171–178). Apart from that, assessing uncertainty and indeterminacy is an essential aspect of dealing with questions of prediction. However, this does not imply that we endorse the anarchic and relativistic “anything goes” perspective that is sometimes evident in post-processual writing. We consider pluralism in the refutationist tradition of setting up multiple models and theories that can be tested and compared in order to reveal weaknesses and thus contribute to the advancement of theory (Bell 1994). In this context, the advancement of theory implies that all interpretations are not inherently equally valid and that, just as we evaluate some sites as more important (to us as archaeologists) than others, the same applies to archaeological explanations and interpretations. Scientific methods are essential to this. From this perspective, quantitative modeling should be taken as seriously as collecting data, developing logical arguments, and producing narratives. From the perspective of predictive modeling, it means that theory building should be an integral part of the effort to predict the location of archaeological resources.

Middle Range Theory

The question then remains: what kind of theory can lead to better predictions, satisfying current theoretical concerns without becoming too complex to handle in

practice? From a modeling perspective, middle range theory is an obvious candidate. Middle range theory was first developed in sociology in the 1950s (see Merton 1968; Raab and Goodyear 1984; although Kosso 1991 claims that the natural sciences were the main source of inspiration for it in archaeology). Merton (1968, p. 52, as cited in Raab and Goodyear 1984) stated the problem with theory like this:

“A large part of what is now described as (...) theory consists of general orientations toward data, suggesting types of variables which theories must somehow take into account, rather than clearly formulated, verifiable statements of relationships between specified variables.”

Crucially, the way in which Merton envisaged middle range theory is by the creation of a logical structure in which working hypotheses can be confirmed or negated and in this way reflect upon the validity of general theory. Middle range theory is then the critical bridge between general theory and empirical data, in which both inductive and deductive perspectives can be effective. Kosso (2006) provides a simple example of how middle range theory is applied in archaeology. Archaeologists interpret off-site scatters of pot sherds as the product of manuring and therefore as evidence for cultivation. The presence of pot sherds that are relatively evenly spread over the area (the empirical data) leads to a logical, deductive, middle-range explanation: the observed pattern can only have come about because people mixed pot sherds with manure and spread this mixture over the land as fertilizer. This implies that the area was under cultivation in a particular period of time, and this can be used as evidence for further claims on the socio-economic situation.

Middle range theory as applied in archaeology was originally methodologically oriented and primarily dealt with site formation processes and the so-called behavioral archaeology (Schiffer 1976; Binford 1977, 1981, pp. 21–30). The focus was upon typically natural processes that could be observed as having a measurable influence upon material remains. Nevertheless, many archaeologists have identified it as an indispensable tool for other types of archaeological research as well (see, e.g., Bogaard 2004; Dobres and Robb 2005). In fact, even post-processualists have routinely relied on it for developing their interpretations of the past (Kosso 1991; Tschauner 1996), and Bell (1994, p. 18) asserts that virtually all archaeological researches are about developing and applying middle range theory. According to Binford (1981, pp. 21–30), good middle range theory should be unambiguous, based on cause and effect rather than simple correlation, applicable to the past by using uniformitarian assumptions, and intellectually independent of general theory. These criteria are however not unproblematic, and we will try to illustrate this in the next section.

Raab and Goodyear (1984) pointed out that most archaeological theories developed by New Archaeologists aimed at both generality and comprehensiveness. In other words, theory was on the one hand supposed to specify law-like assumptions with universal applicability (“covering laws”) and at the same time predict phenomena in considerable detail. The tension between these goals is evident, and when applied in this way it is virtually impossible to connect fundamental theory to actual data. As explained earlier, radical post-processualism takes the other extreme and to a large extent denies the possibility (and therefore the value) of generalization and prediction. Raab and Goodyear instead proposed to

develop theory that is directed to explaining aspects of cultural systems that *can be explored empirically* and that might, at some point, be connected to a higher-level theory. Having written their paper just before the rise of post-processual archaeology, they recommended a form of systems theory that also includes social aspects. As mentioned earlier, examples of this way of thinking such as cognitive archaeology and complexity theory have not met with general acceptance among archaeologists.

Afterthought

Having tried to recapitulate the major currents in archaeological thinking connected to quantitative modeling over the past four decades, we cannot help but look in wonder at the theoretical labyrinth that seems to have been erected. As discussed earlier, other sciences seem to be somewhat less concerned with the theoretical justification of quantitative model building and in that sense may be considered too naïve in their approaches to predictive modeling. On the other hand, the almost impenetrable jungle of archaeological theoretical writing has not helped us very much towards a better integration of quantitative modeling in archaeological research. Theoretical notions are often defined in such an abstract way that it may be quite hard to figure out whether we are dealing with, for example, middle range theory at all.

Between Merton's definition and the interpretations of Raab and Goodyear, Binford, Kosso, and Tschauner, a substantial gray area is found in which these writers are struggling to position middle range theory in archaeological research practice. The same can be said for much post-processual writing. Dobres and Robb (2000), for example, reflect on the lack of a clear definition of agency in archaeology (even citing 12 definitions of it found in literature) in their introduction to 16 more papers on the subject that only seem to agree on the fact that agency should be included in archaeological research in some way.¹ They even admit (Dobres and Robb 2005) that "few ideas so popular in 21st century archaeology have led to such sparse methodological developments." Archaeologists appear to actively employ popular terms such as "agency" without specifying or perhaps even understanding their definitions.

In our attempt to give theory a more prominent place in predictive modeling, we will therefore not aim to solve it all and provide a universal method for dealing with it or to give all interpretations of theoretical schools their due place in it. We are however convinced that there are practical ways to better embed archaeological theory in predictive modeling. In elaborating this, we will mainly stick to the concept of middle range theory as defined by Binford (1981) and Raab and Goodyear (1984), including aspects of agency theory and cognitive archaeology, and acknowledging that multiple models may be necessary to arrive at sound predictions.

¹ The clarification of the subject is not helped by definitions like 'Agency is always situated in the resources of space/time, a being-in-the-world whose actions carry the past into the future and which make reference to absent places in the locations of its own operations' (Barrett 2000). It is just one example of a tendency to mystify aspects of archaeological theory in terms that are almost as incomprehensible to the uninitiated as mathematical equations.

Elucidating Middle Range Theory: Urnfields and Settlement in the Southern Netherlands

As an example, we can take the distribution of Late Bronze Age and Early Iron Age settlement in the southern Netherlands (approx. 1100–500 BC). This particular example is taken because it constitutes an issue of considerable debate in Dutch archaeology, and this debate has clear implications for predictive modeling in the area. For the moment, however, a predictive model based on the various theoretical approaches suggested is still lacking, and the example is only used here to illustrate the complexities of translating various theoretical approaches into middle range theory that could be used for predictive modeling purposes.

Several important cultural changes are observed at the transition of Middle to Late Bronze Age in the area, and these are summarized as follows (Roymans and Kortlang 1999, p. 36):

1. A new mortuary ritual is introduced in the form of urnfields. These are collective cemeteries of small barrows at stable locations, whereas earlier burial mounds were more widely spaced and probably only used by individual family units.
2. New, decorated ceramics are introduced that are particularly abundant in grave contexts.
3. Bronze objects become more common, including prestigious weaponry. This is attributed to a more intensive circulation of these goods through Atlantic and Central European exchange networks.
4. A new system of arable farming is introduced in the form of *celtic field* agriculture,² implying a higher degree of collective organization.

The driving force behind these changes is thought to be the development of more hierarchical forms of social organization and increased competition between groups. Roymans and Kortlang (1999) argue that population increase and corresponding pressure on the land created social problems that were (at least partly) dealt with by changes in the mortuary ritual. Each local community probably consisted of three to six families, controlling a territory comprising a celtic field complex with dispersed farmsteads, an urnfield, and a peripheral zone of uncultivated land used for grazing, collecting wood, *etc.* Each known urnfield is therefore an indication of such a territory, and the farmsteads would move around in the territory with an average life span of 20–30 years. This system contrasts with the preceding Middle Bronze Age, where a greater degree of mobility of settlement is assumed, and the primary level of organization is the family unit rather than the local communities of the Late Bronze Age and Early Iron Age (Gerritsen 2003; Roymans *et al.* 2009).

On the basis of the number of known burial sites, a steadily rising population is assumed during the whole period (see also Gerritsen 2003; this assumption is

² Celtic fields are small (20–40 m in width) rectangular or square plots of land that are separated by low, man-made earthen ridges. They are found all over Northwestern Europe and date from the Late Bronze Age until the Early Middle Ages. Spek *et al.* (2003) distinguish two distinct phases in the development of these field systems (at least in the northern Netherlands). In the first phase, land was cleared and leveled and an agricultural system with a low level of manuring and relatively long fallow periods was established. In the Late Iron Age and Roman Period, the field boundaries were raised and widened to form broad (8–12 m in width) earthen ridges that were intensively cultivated.

however challenged by Fokkens 2002, pp. 143–144). Compared to the Middle Bronze Age, a threefold increase in population is postulated at the beginning of the Early Iron Age. A process of “filling up” of the landscape is assumed by fission of small groups of people from existing settlements. Survey evidence suggests that this filling up was not only done by occupying empty areas but also at the expense of existing territories, thereby decreasing territory sizes and distances between settlements. Similar processes are described by Bintliff (1999) for a number of cases in other regions and time periods. Roymans and Kortlang suggest that control over land therefore became a crucial factor for the longer-term survival of local communities, who developed new social mechanisms to deal with this changing relationship with the land. By means of kinship ideology (marriage rules, patterns of inheritance), the transmission of rights on land could be restricted to the local group. The role of the urnfields in this process is considered crucial: they provided a long-term focus for the community and strengthened collective identity. Furthermore, they may have functioned as territorial markers, explicitly stating the claims of the local community on its territory.

All in all, a coherent interpretation of Late Bronze Age/Early Iron Age settlement in the southern Netherlands is presented. But is it a suitable theory for predictive modeling purposes? Earlier we stated that a form of middle range theory is needed for developing good theory-based predictive models. But do we actually have a good example of middle range theory here in the sense of Binford (1981)?

- First of all, the theory is *unambiguous*. For the development of the territorial system associated with the urnfields, a single and coherent theoretical perspective is offered. Of course, this does not mean that alternative theories might not be available.
- The theory clearly provides plausible *causes* (population increase and associated social change) for demonstrable *effects* (the introduction of celtic field agriculture and a change in mortuary practice coupled to territorial division of the land). The actual interplay of causes and effects however remains somewhat vague.
- The *uniformitarian condition* will be problematic for any attempt to explain social change in prehistory as there are no present-day correlates available that could be used as benchmarks. This particular theory is no exception.
- Lastly, the model is probably not completely *independent of general theories* on human behavior. Binford is not very clear on how to separate general from middle range theory but emphasizes that middle range theory should focus on processes (the core of processual archaeological thinking, after all) and the variables playing a role in these. An explanatory model invoking kinship ideology as a means to exert land control is then too general as it does not specify the elementary processes resulting from this ideology that influence the settlement pattern.

So, the Binfordian demands on good middle range theory are only partly fulfilled. As the uniformitarian condition will be problematic for all theories that try to explain social change, we can conclude that this particular requirement does not invalidate it as middle range theory. Binford seems to have been unnecessarily restrictive in this respect, possibly because he envisaged middle range theory to be used primarily for explaining site formation processes. However, it can still pose problems in practice,

especially since Raab and Goodyear's definition explicitly restricts the development of middle range theory to issues that are open to empirical exploration.

Furthermore, various authors have pointed out that it is very difficult and perhaps impossible to devise middle range theory independently of general theory (see, *e.g.*, Hodder 1991). Bogaard (2004) interprets this condition in the sense that general theory should be developed from various mutually independent, middle range theoretical approaches. In her case study, she used plant ecological theory relating to the behavior of weeds to draw conclusions on Neolithic farming practices in Central Europe. Plant ecological theory is completely independent from the (more usually applied) anthropological theories on crop husbandry regimes and led her to reject the (general) theory of shifting cultivation in favor of intensive garden cultivation. The relation between middle range and general theory is therefore bottom-up and not top-down: plant ecological theory can have implications for more general theories of human behavior, but these theories will not influence plant ecological theory itself. In this way, a framework of layers of theory can be developed, starting from the most elemental building blocks to the ever more complex general theories. The higher layers of general theories will never be independent of the lower levels of theories used to build them, but it should not be the other way around. However, it remains a matter of debate if this form of independence is always attainable and even desirable as justification for any theory will have to come from evidence, but theory should also be used to justify evidence (Kosso 2006).

Most importantly, however, we find a problem in the application of Roymans and Kortlang's theory to predictive modeling in its inadequate specification of how causes and effects might have operated. It does not explain *why* or even *if* population pressure should lead to the social change from a society based on relatively mobile family units to one with more stable local communities. Not all of this is relevant to predictive modeling: the assumption that urnfields may have provided a strengthening of collective memory does not have many consequences for a predictive model, apart from re-affirming the undeniable fact that the urnfields were used for much longer periods than Middle Bronze Age burial sites. However, when considering the origins of celtic field agriculture, we have to be more wary as it is far from clear how the celtic fields were actually used.

In general, a causal link is assumed between population pressure and changes in agricultural production regime. Boserup (1965) considered agricultural intensification (*i.e.*, a reduction of the fallow period, allowing for more harvests from the same land) to be the way prehistoric farmers coped with population pressure. This intensification is primarily made possible through the introduction of technological improvements like the plow, animal traction, and manuring. Population pressure in itself, however, is not sufficient to always explain intensification (see amongst others Carlstein 1982; Ellen 1982; Barker 1985; de Hingh 2000; Thurston 2007), and the danger of circular reasoning is obvious as it will be hard to prove whether the adoption of new farming techniques preceded or followed population growth. Furthermore, extensification (taking more land into production while maintaining the same farming methods) might be an equally well-suited strategy to cope with an increasing demand for food. Examples of intensification of agriculture without apparent population pressure are known as well, like the installation of a system of surplus production for elites or producing for a market-based economy (Thurston

2007). The issue is obviously highly relevant to predictive modeling: in a model with extensive agriculture, evidence of farming could be found in a relatively large zone around the settlement, and settlement territories would need to be relatively large in comparison to population size. An intensive agricultural system implies a much smaller zone of farming activities and, ultimately, a potentially higher density of settlement.

From Roymans and Kortlang's writings, it is not immediately clear whether they see celtic field agriculture as a more intensive system than the itinerant farming practices of the Middle Bronze Age. Roymans and Theuvs (1999) specifically mention celtic field agriculture as a form of intensification; this is however a relative qualification. Compared to the Neolithic period, celtic field agriculture is certainly more intensive, but compared to the Late Iron Age and Roman period it is extensive (Spek *et al.* 2003; Gerritsen 2003). De Hingh (2000, pp. 210–211) concludes on the basis of palaeobotanical research that relatively intensive agriculture with manuring was already practiced in the Middle Bronze Age. This would seem to indicate that the celtic fields themselves are not evidence of further intensification, especially since no technological innovations are suggested that might have allowed for higher agricultural production. Instead, a diversification in crops is observed in the Late Bronze Age. Population pressure may have played a role in this as diversification can be a means to reduce the risk of crop failure. The celtic field system itself however primarily seems to reflect a change in land ownership and land management, perhaps related to the broader range of crops grown and the more intensive care these needed. Possibly, fences were erected or hedges planted to keep cattle in the fields in fallow years.

So, how should we translate all these into a predictive model? The central issue is how these developments influenced the spatial patterning of urnfields, farmsteads, and celtic fields as these are the primary locations where archaeological remains may be found.

It seems logical to assume that the first settlements were installed before the urnfields, so if we can get a grip on settlement location factors then we will have a good idea of potential urnfield locations as well. The primary factor influencing settlement location choice will then have been the availability of sufficient arable land to feed a local community of three to six families (20–40 people; Roymans and Gerritsen 2002). Settlement location preferences will in that sense not have been very different from the earlier Middle Bronze Age period. Given the debate on the origins of agricultural intensification, we might either assume that these larger communities needed a larger territory per settlement than their predecessors (in which case small, isolated pockets of suitable land will no longer have been preferred) or that they took the option of intensification in order to feed more people from a similarly sized territory. Fokkens (1998) suggested that a territory of ca. 300 ha per urnfield would be sufficient to feed the population, including grazing land. Once such a location was found and land was taken into agricultural production, the settlement's territory will have become more or less fixed, with an urnfield probably installed soon afterwards in the center or at the edge of the celtic field system. What happens next is crucial to the model: either we assume that the process of fission and filling up of the landscape with settlements went at the expense of an existing territory, in which case agriculture could only have kept up

with food demands by means of (further) intensification, or we assume that the groups that split off from existing settlements went to non-farmed areas, and intensification was not necessary.

All in all, it seems that even with conservative estimates of the area necessary for agriculture and husbandry there is little reason to expect problems with population pressure in the early stages of the urnfield period, and a model in which groups would split off from existing settlements to colonize a new territory seems to be the most appropriate (see also Gerritsen 2003). In the later stages, when most of the available arable land was taken into production, intensification may have become more important, especially since it assumed that the poorer soils in the area became depleted (Roymans and Gerritsen 2002). And, in fact, both processes may have operated simultaneously in different parts of the region. We might therefore try to apply both models to this specific situation and test these. This brings us back to the core of the processual/post-processual debate. An important reason for early post-processualists to reject Binford's concept of middle range theory (Hodder 1991) is the fact that the processes involved in archaeological site formation are inherently inaccessible. The uniformitarian assumption can therefore never be fulfilled, and the theory's assumptions can never be tested empirically. We may be restating the obvious here but we want to bring the point home once more: predictive modeling allows us to specify probable outcomes of theoretical notions, even if the processes themselves are inaccessible, and these outcomes can perfectly be tested using independent observations (see also Kosso 1991).

Cognitive Predictive Modeling

In the preceding section, we have not yet tried to answer the question of how to implement archaeological theory in predictive modeling. We contend that most, if not all, issues pertaining to the geographical distribution of archaeological remains can be approached by applying a combination of the principles of cognitive archaeology and the analysis of post-depositional processes. In this study, we will not go into the latter issue but instead concentrate on the cognitive aspects of human spatial decision-making. We have to keep in mind that, in predictive modeling, we are trying to model a rather specific, testable implication of archaeological theory, *i.e.*, the influence of human "behavior" (activities, practices) on the accumulation of the material record at the regional scale. For this, we need a modeling structure that can translate the relevant aspects of human behavior into spatio-temporal terms. The basics of such a method are explained by Whitley (2004, 2005) and will be shortly recapitulated here.

The nature of human decision-making has been a topic in many areas of research including psychology (*e.g.*, Shanks *et al.* 1996), computer science (*e.g.*, Oliver and Smith 1990), and philosophy (*e.g.*, Hume 1739; Hitchcock 1996). Research into the cognitive basis of spatial patterning, including the specific interest in location modeling or cognitive mapping, has long been the domain of human geography (*e.g.*, Downs and Stea 1977; Tobler 1993), economics (*e.g.*, Weber 1929), sociology (*e.g.*, Christaller 1935), and even linguistics (*e.g.*, Levinson 1992). The precise nature of explanation, causality, and probability has been addressed by numerous researchers in the philosophy of science (*e.g.*, Popper 1959; Nagel 1961; Hempel

1965; Salmon 1998), mathematics (e.g., Pearl 2000), statistics (e.g., Cox 1992), and computer science (e.g., Besnard and Hanks 1995) to name just a few. With respect to the principal influences upon the theoretical perspective taken in this section and the development of the model provided here as an example, the human geography literature (e.g., Cohen 1985; Downs and Stea 1973, 1977; Gärling and Evans 1991; Ittleson 1973; Kitchin and Freundschuh 2000; Kitchin and Blades 2002; Moore 1979; Saarinen *et al.* 1984) is extremely important. The inclusion of structures and frameworks from probability theory, cognitive archaeology, and causality (e.g., Pearl 2000; Renfrew and Zubrow 1994; Salmon 1971, 1998) are integral to this example as well. Likewise, the enormous influence of recent advances in agent-based modeling for social and socio-natural systems, particularly the three edited volumes by Kohler and van der Leeuw (2007), Miller and Page (2007), and Epstein (2006), cannot be understated. This is, of course, supplemented by the vast predictive modeling literature already cited in this article, as well as biological and economic modeling studies—specifically the archaeological applications of optimal foraging theory (MacArthur and Pianka 1966; Emlen 1966; Bettinger 1980; Stephens and Krebs 1986; Simms 1987; Kelly 1995; Winterhalder and Kennett 2006), the diet–breadth model (Hames and Vickers 1982; O’Connell and Hawkes 1984; Winterhalder 1987; Smith 1991; Grayson and Delpech 1998), central place foraging (Orians and Pearson 1979; Stephens and Krebs 1986; Jones and Madsen 1989; Metcalfe and Barlow 1992; Bettinger *et al.* 1997; Bird 1997; Grayson and Cannon 1999; Zeanah 1996; Thomas 2008), and—more distantly—prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992; and Wakker *et al.* 2003). The discussion which follows can be seen as being based on an amalgamation of the views, structures, debates, issues, and processes originally presented by these sources.

All predictive models have one thing in common: they are expressions of a probabilistic relationship between human behavior and prior existing spatial conditions. Correlative (inductive) models assume that the current distribution of archaeological sites is a direct reflection of once observable spatial characteristics that were consciously and/or subconsciously selected as locations for human behavior. Statistically assessing the relationships between known sites and such characteristics leads to a predictive formula. The key elements of that formula are the currently measurable variables which represent the *initial conditions* of the past and the actual products of behavior.

Such a predictive formula, though, lacks two key elements: causality and cognition. Obviously, correlative modelers are not assuming that there is no human agency or decision-making in site selection. Yet, the nature of the reductionist statistical analysis used in such models precludes any attempt to understand what cognitive steps are actually present in the process. In between the assessment of the initial conditions and the selection of the location for a (typically unspecified) behavior is an assumption that some cognitive process took place which led to the behavior which caused the deposition of the site. Absence of a site in a correlative model implies the absence of behavior or the presence of behavior of no importance to archaeologists. This makes the assumption that we are concerned *only* with the physical products of the archaeological record itself.

By empirically showing the correlations between sites (or non-site localities) and the modern proxies for past initial conditions, there is presumed to be no need for

understanding the in-between step of cognition at all. Thus, correlative models place this process in what is essentially a “black box.” Cognition is implied but is of no direct consequence for identifying correlations. In addition, no attempt is made to understand the relationship between behavior and the deposition of artifacts or features; it is assumed that there is a direct and consistent correlation. The mechanics of this “black box” are, in essence, the realm of middle range theory.

These mechanisms can be expressed as units which are causally related to each other (Fig. 4). Each of its components can be classified as belonging to one of three different categories:

- *Conditions*—those characteristics which can be (or were at one time) observed in the landscape (or more properly the land unit under evaluation)
- *Events*—a behavior or action bounded in space–time which is typically not directly observable today but which may have had physical consequences (conditions) that are at least partially observable
- *Decisions*—cognitive choices which took place at one time in the past (or are made today) which bridge the gap between observations of conditions and the activity of events.

Decision-making is dependent on the acquisition of information. This can be either *direct* (as in something perceived with the immediate senses: vision, hearing,

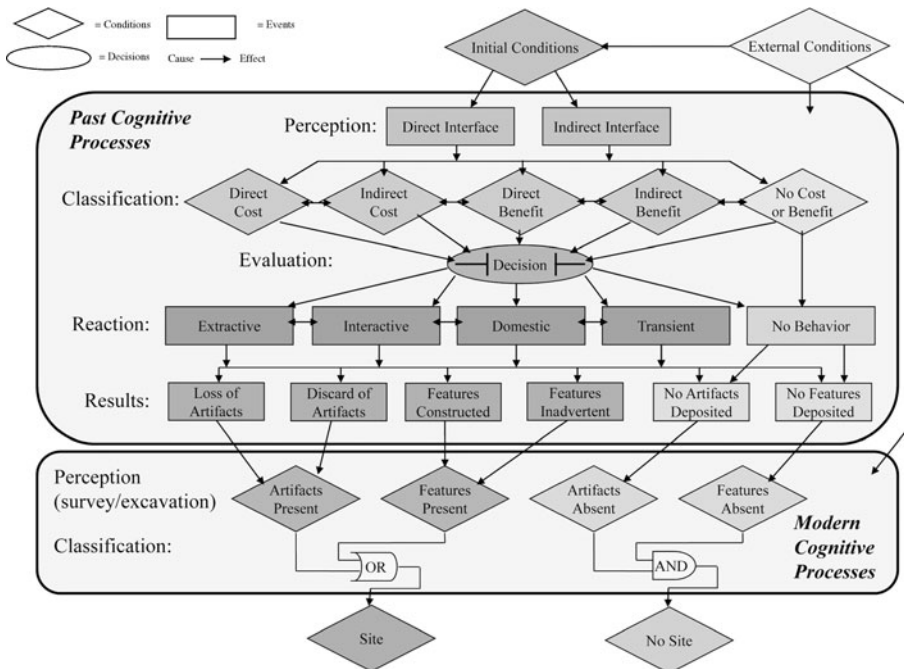


Fig. 4 An outline of the cognitive predictive modeling approach. Decision-making is thought to be the result of conscious or unconscious weighing of the perceived costs and benefits of a particular action, given certain initial conditions. Decisions taken may result in a form of behavior that leaves a material imprint. After Whitley (2005)

taste, touch, and smell) or *indirect* (information gained from other sources: second-hand description, previous experience, or pure speculation). All the direct and indirect acquisition and processing of information represent the cognitive process of *perception* (or acquiring *knowledge*).

After perception, each initial condition is classified (either consciously or subconsciously) with regard to its beneficial or adverse effects on site selection, *i.e.*, a cost–benefits analysis. Even though something may have no direct cost or benefit on the behavior, it may affect how other direct or indirect variables are perceived and classified. Once conditions are classified, a decision is made regarding whether to carry out some behavior. That determination is not based on one cause-and-effect relationship with a single variable. It is, in effect, a decision which is triggered only when sufficient information has been processed through evaluating a number of influential variables, presumably all of those which have a direct or indirect cost or benefit for the behavior. In actuality, it is likely that many behavioral decisions were made on the basis of partial or incomplete information.

This process does not imply that people are always fully aware of the rationale or reasoning behind spatial behaviors. Many behaviors were probably encoded into the neural pathways of past people in such a way that the cognitive processes of site selection were carried out as a matter of subconscious immediacy, based on having learned to heuristically assess a spatial environment for key indicators rather than to fully evaluate their surroundings every time. Clearly, this becomes easier and more predictable for commonly recurring behaviors, and full evaluation of all the possible costs and benefits probably only occurs for uncommon or unfamiliar behaviors. Whether the evaluation is conscious or subconscious though, the structure of the process is still the same.

In the final probability assessment, it is this very decision which is being modeled. A probability formula should, then, be as explanatory and as expansive as possible and not a reductionist *lowest common denominator* of suitability for all behaviors or events. Certain kinds of behavioral event are more likely to produce the intentional discard of artifacts, the unintentional loss of them, and/or the intentional or unintentional creation of features. But artifacts and features are not uniformly representative of past behaviors. In fact, some highly significant spatial behaviors result in the deposition of no artifacts or features whatsoever. Likewise, the preservation of certain kinds of material strongly affects how we see archaeological sites as representative of behaviors. One of the key elements of cognitive predictive modeling is the recognition that interesting and important cognitive processes may result in the use of land units, or vast landscapes, without any archaeological component being deposited. This must also be distinguished from the avoidance of areas for entirely different reasons.

This brings us to the second layer of cognition within the site selection framework: the series of cognitive processes carried out by archaeologists. Past cognition does not result in “sites” *per se*, rather in spatially dispersed material results which unevenly represent many incidences and several broad categories of past human behavior. Our modern cognitive process of perceiving these material by-products (through survey and excavation) and classifying them into meaningful clusters (*i.e.*, sites) completes the causal chain from the initial conditions to our typical units of study.

Example: the Georgia Coast Model

Ultimately, though, how do we turn the cognitive modeling approach into a practical causal explanatory predictive model? When moving from model structure to actual application, we have to take care how to quantify the cost–benefit evaluations. Most desirable would be to translate the costs and benefits into similar “spatial currencies,” so to speak. This can be done relatively easily in cases dealing with subsistence activities, where it is basically a matter of specifying “energy budgets.” The amount of caloric energy that can be provided by a particular spatial unit (in the form of wild animals, cereals, or any other foodstuff) can then be compared to the amount of energy needed to actually collect the calories (including the energy required to maintain the resource, to develop tools and techniques to collect it, to travel to the site of its collection, to harvest it, to process it, consume it, store it for future consumption, and to dispose of it). Several studies in the 1970s and 1980s linked caloric travel costs to site catchment analysis (e.g., Ericson and Goldstein 1980; Styles 1981; Gregg 1988) and are still very useful for spatial modeling of (potential) subsistence activities (e.g., van Hove 2003; Goodchild 2006). The approach can be used for general predictive purposes by modeling the overall difficulty of terrain access and mapping the zones where travel will be least prohibitive and therefore the overall benefit of settling with regard to a specific subsistence activity is highest. Although clearly limited in their ability to quantify non-travel cost issues, these models were some of the first to link cognition and calories in a spatial context. The following discussion provides an example of a more complex cognitive model (Whitley *et al.* 2009) designed largely for explanatory purposes but which has an inherent predictive capacity that illustrates many of the theoretical ideas outlined above.

The Lower Coastal Plain of Georgia (Fig. 5) is a flat, wet, and heavily forested area. Along the coast, we find estuaries protected by barrier islands. Behind these islands lie vast expanses of salt marsh with shallow muddy tidal flats and emergent grasses. At the seaside of the barrier islands, long stretches of narrow sandy beaches backed by dunes are found. The ends of the islands give way to the fast-moving and variable tides at the mouths of wide slow rivers which traveled several hundred kilometers from the Piedmont to the Atlantic. The soil types in the area are all very similar. They are generally only moderately suited for agriculture, poorly drained in low elevations, and excessively or well drained in the slightly higher ones. Today, modern logging has changed most of the upland climax growth forest to a denser scrub understory with mixed evergreen and deciduous forest, but the marshlands remain much as they were at the time of the first European contact.

There are currently nearly 6,000 archaeological sites recorded within the terrestrial portion of this area, representing more than 10,000 years of occupation. The environment however is not very conducive to building a correlative archaeological predictive model using the standard available set of “environmental” variables. The general absence of steep slopes and the ubiquitous presence of freshwater make it impossible to use those variables as a means to limit the expected distribution of settlement choice. Soil type also does not limit site selection because the archaeological sites from all periods are known from virtually all soils.

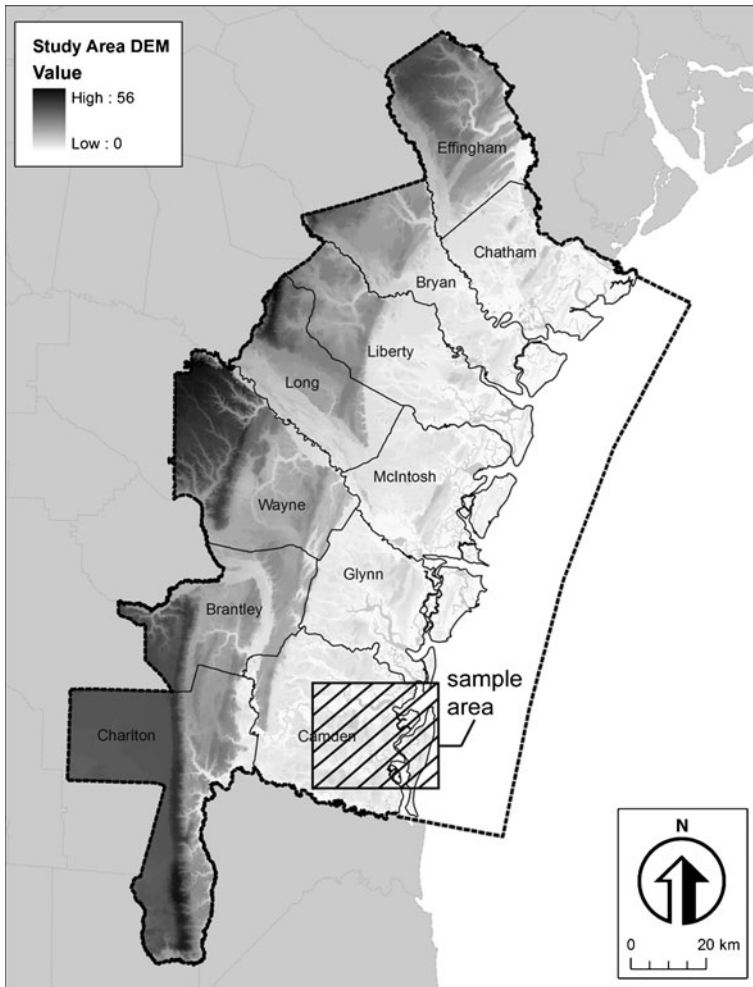


Fig. 5 Location of the Georgia Coast study area

The study area covers 4,669,484 acres (both land and water), and the overall paleoeconomic model exists as three distinct elements: (1) the Habitat Model (HM)—an interpretation of the intensity of the correlation between a given forage category and any map unit for each month of the year, based on the strength and distribution of key elements (*i.e.*, attractors) in their habitat; (2) the Available Caloric Model (ACM)—an interpretation of the total amount of calories that could be expected in any map unit, given a habitat model value and population density estimates for each forage category for each month of the year; and (3) the Returned Caloric Model (RCM)—an interpretation of the amount of calories for each forage category, each month, that could be extracted from any map unit, given the available calories, the technological limitations, and the costs of acquiring, processing, and transporting the targets.

The habitat model is defined by 37 forage categories (derived mostly from Thomas 2008, but expanded); some of which are individual species (such as “white-

tailed deer”) while others are combinations of numerous species based on family or genus groupings (such as “freshwater turtles”) or size/habitat limitations (such as “large saltwater fish”). These groupings include both wild faunal and floral resources, as well as domesticated (or semi-domesticated) species. Ultimately, weighted additive formulas were developed for each forage category using 15 baseline environmental attractors (derived from the USFWS National Wetlands Inventory maps, the NRCS Soils Survey Geographic Database, the USGS National Elevation Dataset, the NOAA Hydrographic Survey Dataset, the GDNR Land Use/Land Cover data, and the USEPA Level 4 Ecoregions Data). The strength of the attractors (*i.e.*, formula weights) were based on both qualitative and quantitative assessments of preference derived from sources such as Thomas (2008), Smith (1992), Reitz *et al.* (2010), NARSAL (2010), Georgia DNR (2010), and nearby states’ online resources giving population estimates or biological statistics. They were broken out by month of the year as well. The results are a series of 444 individual GIS surfaces (37 forage categories * 12 months), each of which covers all 4,669,484 acres with a resolution of 30 m (900 sq m). Each of the 20,996,530 datapoints (referred to here as a “map unit”) has a decimal value ranging between 0 and 1, which represents the predicted habitat suitability for each forage category for each month for that 900-sq-m location.

To transform the habitat model into a caloric expression (the Available Caloric Model), a series of variables were generated, including the population density of each forage category (per 30 m map unit), their average size, the ranges in weight by gender, the adult/juveniles and gender ratios, the usable calories per kg, the average number of offspring, the mean birth months, time until sexual maturation, prime harvest months, crop yield, and fallow cycle. All of these assessments were based on the same sources of information as above or were projected as reasonable quantitative estimates where no specific data, or only qualitative data, were available. It should be noted that the framework of this GIS model is not dependent upon the initial values chosen and inserting better or alternate data is always possible when it becomes available.

The analysis of these variables allows a projection of the expected number of calories per 30 m map unit of prime habitat (*i.e.*, HM value of 1), as well as their monthly resilience (a function of the estimated population density, their rate of survival assuming a stable population, the length of time it takes for an individual to reach reproductive maturity, and their mean number of offspring). The maximum available calories were calculated by forage category and month, assuming climatic conditions similar to today and a stable (modern) sea level since 4500 BP. The caloric values were then multiplied by each of the appropriate HM surfaces. The result is the transformation of each of the 444 surfaces into an expression of the ACM. The ACM is, in essence, a representation of the exploitable ecological landscape by species and month.

The Returned Caloric Model (RCM) is built upon estimates of the bracketed minimum and maximum expenditure of calories per day, per individual, in maintaining, collecting, and processing each of the forage categories. Thomas (2008) provides some of the estimates of collection and processing time for many of the species involved, while others are derived from Smith (1992) or other sources. The model assumes that maintenance (*e.g.*, tool manufacture, making nets and weirs,

habitat improvements, *etc.*) and collection costs (*e.g.*, search times, setting and checking traps, *etc.*) decrease over time (between 4500 and 300 BP—the applicable time frame), while hunting success and harvest rates, as well as storage potentials, increase with certain specific technological changes (*e.g.*, transition from spear to dart to bow and arrow, development of pottery for storage, increased trade for high-quality lithic materials, *etc.*). The return rates of calories expended for calories collected are thus calculated as a percentage and applied to the ACM.

The RCM is intended to be used interpretively for locales where we already have a reasonable understanding of settlements and potential populations. With that knowledge in hand, we can generate travel costs based on cost-distance evaluations using both foot and canoe caloric friction and use it to calculate potential caloric surpluses or deficits for a given site, at a given point in time, with regard to a given season or month and based on projected populations and dietary preferences. For areas where we do not have extensive prior knowledge of site locations, we would generate a summary probability surface based on a composite model. One example of such a model (dietary preference) was devised for the period between 4500 BP and 300 AD using all 37 forage categories. It is pinned at either end by proportional estimates based on Late Archaic and Contact Period faunal and floral assemblages. The values in between are calculated as exponential or logarithmic percentages of the difference between the end values. The dietary model is not intended to definitively represent an archaeological interpretation of the range in past diet; rather, it is meant to provide a simulation which can be used to express an overall impression of the transition from Late Archaic to Contact Period diet as we currently understand it. That allows us to use those proportional estimates to weight the returned calories by forage category and month and combine them into a composite probability formula for any point in time. This has not yet been applied to the entire study area; however, another example was previously provided by looking at one detailed aspect of the study area.

An area near the boundary between Georgia and Florida measuring approximately 36 by 28 km was presented as a sample area (see Fig. 5) within which to produce and test a cognitive predictive model (Whitley *et al.* 2009). A simple deductive predictive model was made using a composite of all available calories by season (such as shown in Fig. 6) and the highest caloric return areas were compared to the known locations of existing sites. In this case, 308 sites fall within the study area. Merely splitting the composite caloric returns surface equally into three categories (low, moderate, and high) produced a Kvamme's gain statistic³ (Kvamme 1988) in excess of 0.80, whereas the best correlative or intuitive models did not produce better gains than 0.57 (Fig. 7). Furthermore, almost all of the sites in low-potential areas were found along the edge of high or moderate probability areas. Their

³ Gain is calculated as follows:

$$G = 1 - \frac{p_a}{p_s}$$

where

- p_a The area proportion of the zone of interest (usually the zone of high probability)
 p_s The proportion of sites found in the zone of interest.

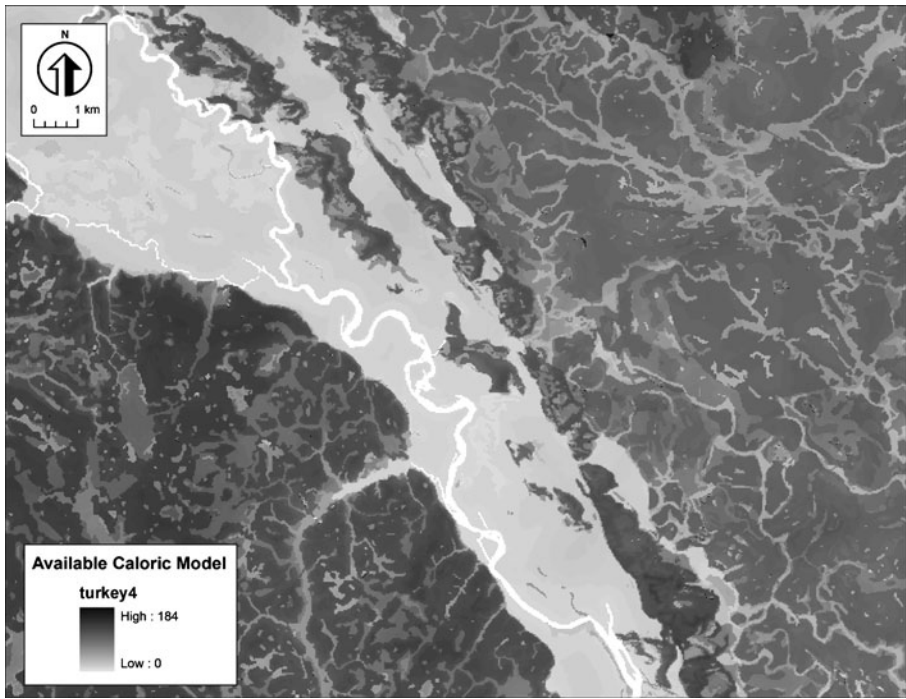


Fig. 6 Expected caloric returns for turkey

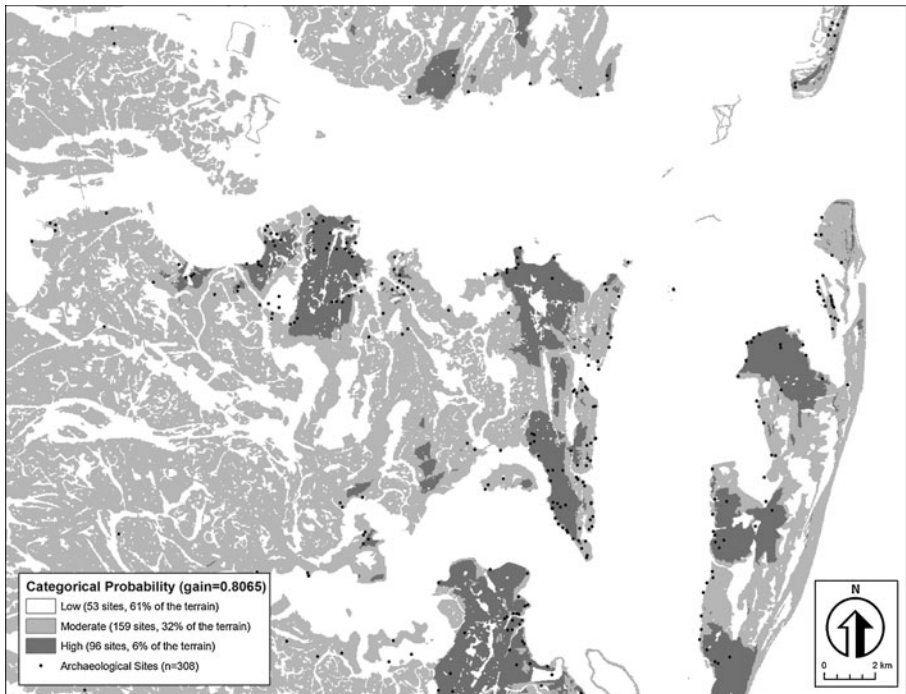


Fig. 7 Predictive map based on the caloric returns model

location can likely be attributed to lack of resolution in the data or lack of accuracy in recording rather than to a preference for low probability areas for some other reason. This was a clear example of the power of caloric surfaces to project potential site locations even with a minimum of interpretation injected.

But developing a cognitive predictive model also means that we try to understand things such as what kinds of activity occurred in which areas, what the nature of resource competition was like, how foraging differed between genders, and how the perceptions of the agents affected their knowledge of resources and their costs for acquisition. This is the middle range theory portion of the analysis. For example, given that some species return only a modest number of calories, while others much more so, the cost-distance (*i.e.*, foraging radius) at which it is no longer efficient to collect them would vary depending on their probable caloric return, ability to be gathered in quantity, dietary attractiveness, potential for other uses, and the current caloric or nutritional stress of the foragers. Similarly, the cost-distance radius at which it is more efficient to process the collected resource rather than bring it back whole (*i.e.*, processing radius) would also be a function of its weight and its processing time or difficulty. Thomas (2008) provides a very detailed discussion of the probable foraging and processing radii for many species or forage categories in the Coastal Georgia region. In general, he finds that for most species an effective one-way-daily foraging radius of 450 kcal consumed (or around 5 km in his estimation) is likely (Thomas 2008, Fig. 11.12). He also charts processing radii as a function of distance and categorical thresholds (Thomas 2008, Table 10.7).

For the Coastal Georgia analysis, caloric distances were calculated from known archaeological locations in the study area (based on 2010 information from the Georgia State Site Files) for both foot and canoe travel. The resulting surfaces can be used in several different ways. First, foraging and processing thresholds can be generated for any species of interest. For example, the processing radii provided by Thomas (2008, Table 10.7) can actually be mapped as buffers around a central place, indicating the point beyond which it is more efficient to process the resource before bringing it back. This helps illuminate areas where we would expect to find processing sites related to central place occupations. This is particularly pertinent with regard to shellfish processing *vs.* consumption and is an example of predictive modeling in a local context.

Assuming an average effective daily foraging radius of 450 kcal, the “catchment areas” around any given location can be mapped (based on foot and canoe travel—*i.e.*, terrestrial and marine foraging, respectively—Figs. 8 and 9 illustrate both for a portion of the study area) and the mean and total caloric returns can be calculated for any given month for any one species or for weighted combinations that represent projections of complete diets. Since this can be done with high-scoring probability areas as well as known sites, it could be used to estimate carrying capacities and populations even in unsurveyed areas.

Because these models are ultimately representations of the spatial distribution of potential energy sources, it may be possible to develop a wide range of interpretive surfaces that consider more complex economic issues (a few examples are illustrated in Fig. 10). The foraging radii (including both proximity- and perception-related fall-off rates) can provide estimates of carrying capacity, relative monthly bounties by forage categories, intensity of resource competition, and surplus and exchange



Fig. 8 Single day's foraging limits for hunting, Late Mississippian/Contact Period

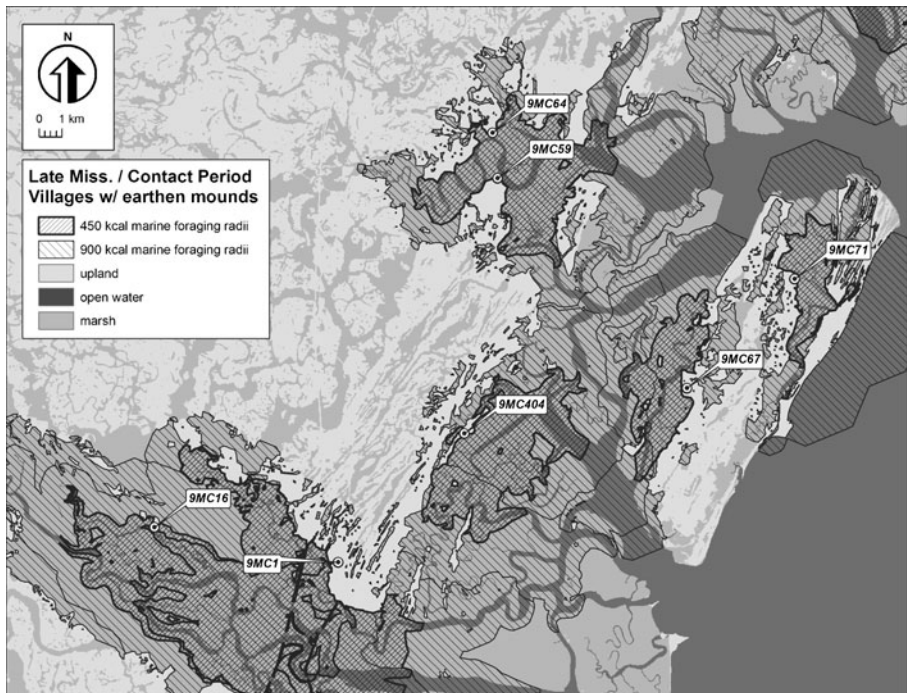


Fig. 9 Single day's foraging limits for fishing, Late Mississippian/Contact Period

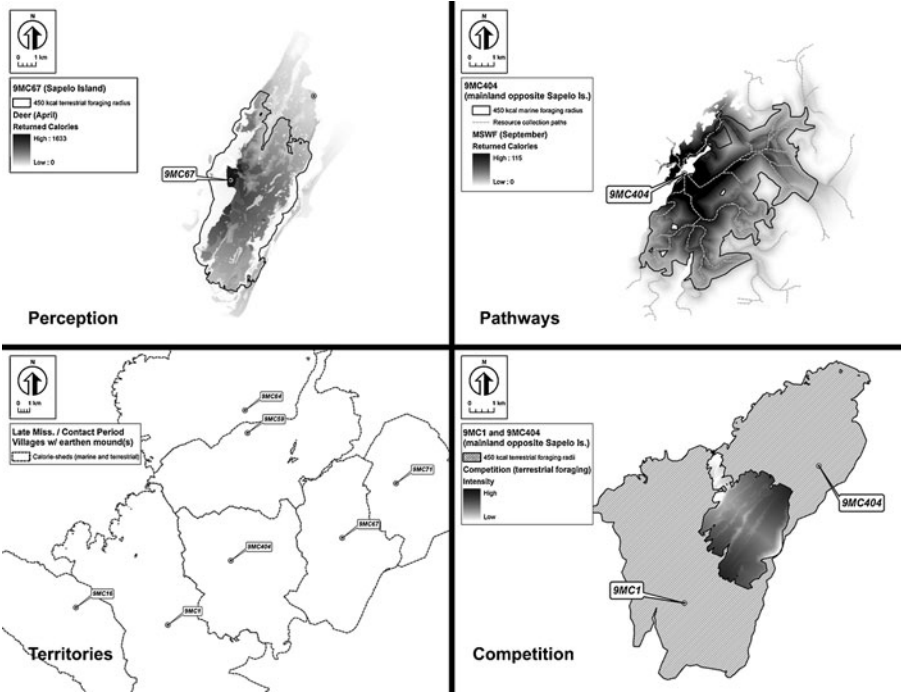


Fig. 10 Illustration of other potential interpretative surfaces resulting from the modeling

potential. Transforming the accumulated caloric distances into pseudo-topographic surfaces can help model and understand the success of resource acquisition pathways (e.g., by performing pseudo-hydrologic or least cost path analyses). Caloric distance can also function as a proxy representation of territories (using caloric thresholds—“calorie-sheds”—as boundaries) or social dominance (larger population centers may occupy a visible “caloric sink” and they might extract calories from adjacent areas through controlling exchange routes or tribute).

With regard to the potential for predictive models, the modeled pathways from very-high-scoring areas would be particularly interesting. These would be the corridors along which resource gatherers would routinely spend a great deal of their time. We would expect that sites resulting from the loss or discard of artifacts associated with daily activities would occur along these pathways, and a predictive model could be generated to capture them.

This example is still a work in progress. The specific surfaces illustrating the available caloric returns and the associated spatial decisions are guided by the environmental conditions, the species present, the acquisition (e.g., hunting, gathering, agriculture) and storage technologies available, and furthermore in the nature of trade, tribute, and the speculation on future returns that may have been part of social contracts. But ultimately these ideas fall within the “black box” of cognition and are in the realm of middle range theory. With respect to Binford’s criteria:

1. Cognitive modeling in this example is unambiguous. The theory presented is straightforward and based on solid principles of energy capture and conserva-

- tion, without being (environmentally) deterministic. Although prediction is implied rather than specifically addressed at this juncture, the basis for making location predictions includes modeling agent-based decisions and perspectives, as well as gender and social differences.
2. The process of decision-making presented is distinctly cause and effect, not correlative. Though correlative measures can be used to test the validity of identified associations, those associations are not presumed upon the existence of the correlations.
 3. Binford's notion of uniformitarian assumptions is satisfied if we consider his meaning to have been that the explanations presented in this model do not require the adoption of general rules of human behavior that do not apply in other areas under other conditions. The only general rule or assumption required by the Georgia Coast Model is that of energy conservation as a primary motivating factor for certain human spatial behaviors.
 4. The model can be considered independent from general theory (in the sense of Bogaard 2004) because it is independent from our assumptions about Native American subsistence practices on the Georgia Coast and their knowledge of the landscape. The theoretical framework itself does not change if we were suddenly to discover that the Georgia Coast natives hunted a different range of animals, grew different crops, or used a different means of transport. Therefore, it can be transplanted to different regions and time periods as well. Only the variables need change, not the theory.

However, in its current form, the model cannot be used directly for aspects of human behavior that are not related to (daily) subsistence activities. The model assumes that spatial units can represent a quantified "value" to the users of the available resources—in other words, a "spatial currency" to be withdrawn, banked, or traded to others. The decisions made reflecting their use or storage are tempered by the costs and benefits involved. However, not everything can be so easily budgeted and evaluated for costs and benefits in terms of caloric returns. In many cases, calories are expended in exchange for ephemeral or often unmeasurable benefits. Energy might, for example, be "converted" into valuable resources in the case of extraction of lithics, metals, or other natural resources, which can in turn be used for other purposes or traded for other goods. It will however, in many cases, be difficult to specify what the currency at the other end of the equation is, let alone the conversion rate.

Take for instance burials. While we can probably specify the energy input needed to go to a particular place and perform the burials and associated rituals, there is no quantifiable output involved. Possible benefits could be of a practical (deposing of the deceased's body), social (strengthening community bonds), and ritual nature (providing a place of access to the world of the ancestors), and all of these have no clear quantifiable or even spatial characteristics, leaving us with considerable problems when trying to predict the most probable location for a burial site. A number of other examples could be given in which the *perception* of costs and benefits, rather than the exact measures of these, will determine the outcome of the evaluation. Evaluating and using a cognitive cost–benefit approach will therefore almost never be a clear-cut case—which is where the multiple modeling framework

comes in. By specifying more than one option and looking at the issue from multiple perspectives, we might be able to come up not so much with the final answer but at least with the most probable one (or a few probable ones).

Many of the earlier attempts to set up and test multiple models were basically constrained by the technical difficulties involved. Even in the late 1990s, setting up and calculating a suite of scenarios in a GIS could take weeks or months, which would make it highly impractical to pursue a multitude of models. While we still do not have software that is particularly geared towards these kinds of exercises, increased computing power has now made it possible to run hundreds or thousands of models in a relatively short time. The modeling itself is no longer the problem—but developing and translating the necessary theory still is.

Some might argue that the modeling procedure suggested does not provide much innovation. And in fact our approach should be seen more as a logical extension of agency theory to older concepts like site catchment analysis and optimal foraging theory than as an inclusion of the “individual emotional, sensational, and experiential aspects that are unique to that individual” (Dornan 2002, p. 322). It is in that sense still closely connected to the systemic, processual way of thinking and might therefore by some be considered fundamentally structuralist and therefore “de-humanized.” However, we hope to have made clear that the approach sketched is not at all limited to ecological or economic systems and is a long way removed from the inductive and intuitive procedures prevalent in predictive modeling. The cognitive modeling framework has already been successfully applied to a variety of case studies, including American slave societies (Whitley 2003), an Egyptian necropolis (Burns *et al.* 2008), and long term hunter–gatherer subsistence dynamics in the drowning Mesolithic landscape in the Dutch province of Flevoland (Peeters 2007), and as far as we can judge it delivers predictions more cost-effectively than (especially) inductive models. It includes agency in a straightforward way by specifying what Cowgill (2000) describes as “recurrent types of context” within which people react with reasoning in regard to their perceived interests. These contexts are specific configurations of social and environmental variables that not so much determine the issues about which people reason but at least narrow these down to a range of what is likely. The key problem to solve is then trying to recognize what might have been the perceived interests of people within their socio-environmental context.

Model Testing

Having established a framework for building theoretically informed, cognitive predictive models, we now want to turn our attention to the other side of the equation: how do we decide whether the model is any good or better than a competing model? Testability (or refutability) is the hallmark of any good scientific theory (Bell 1994), and, as we stated earlier, modeling is a way of opening up a theory to testing. However, we have to be careful with the terminology used with regard to model testing. Predictive models are often supposed to be “validated” by confronting them with (measured) data. Note that, in the case of predictive models as they are currently applied, the only testable implication is found in the prediction of

the geographical concentration of archaeological remains, in particular settlement sites. Repeated consistency between model output and measured response is then taken as proof that the model is validated or even verified. However, in its true sense, a valid model is one that is without logical errors and internally consistent (Oreskes *et al.* 1994). And in order to be verified, the model should be “true,” *i.e.*, produce accurate results under all circumstances. It is not hard to see that most sciences will have tremendous difficulties in producing these kinds of model results, and successful examples are restricted to things like the prediction of celestial mechanics, for example. In all other cases, confirming observations do not prove the veracity of a model; they only support its probability.

Predictive model testing is therefore a procedure with two aspects. First, we have to establish the internal consistency of the model, and we could call this the validation phase. Most importantly, we should be looking for possible conceptual errors in the model (and thus in the theory behind the model; Bell (1994) refers to this as consistency testing). In this stage, factual error should be considered as well: have we used data sets that correctly represent the parameters that we want to include in the model? This relates both to the “environmental” factors (like using an accurate digital elevation model) as well as the archaeological data whose chronological and interpretative accuracy should be assessed.

In predictive modeling, testing is either done through peer review, statistical evaluation, or a combination of both, using new survey or excavation data when they become available. Peer review (or expert judgment) is currently the only method used for identifying conceptual error in the model and the underlying theory, and more formalized methods for doing so seem to be lacking anyway. Statistical evaluation of the model results (empirical testing) will not suffice for this as it can only indicate that there is something wrong with the model predictions but can never tell us directly what the cause of the problem is: conceptual error (interpretive mistakes) and/or observational bias and/or factual error. In theory, by removing observational bias, we might be able to ascertain that unsatisfactory predictions point to conceptual and/or factual model errors. Spotting and “repairing” conceptual error could be handled by applying different theoretical frameworks to the same situation to see which framework offers the best predictions. This would seem similar to calibrating a model, a procedure in which the interactions between the model components are varied in order to tune model output to conform to known observations. Calibration, however, while standard practice in statistical (regression) modeling, ignores all problems of conceptual and factual error as well as observational bias.

In the second stage of testing, the model should be confronted with observational data to establish its predictive power: how good are the predictions and what is the uncertainty involved? We could call this the evaluation phase (Oreskes 1998) where both positive and negative results are possible, and these can be used to establish whether the model is good enough for its purposes. The evaluation phase basically legitimizes the use of the model, *e.g.*, for management purposes. This somewhat contrasts with Bell’s (1994) idea of empirical testing. In his (refutationist) view, testing can only serve to disprove theories, but the data themselves will never do this. It is the scientist who *decides* that data refute a particular theory. Obviously, a model may be rejected after realizing that the available data contradict its outcome,

but the actual criteria on when to refute a model are unspecified and using statistical significance levels is only partly helpful in this. By defining probability ranges as “acceptable” outcomes, anomalies may be explained away as unimportant: it all depends on the willingness of the scientist to accept anomalies as being significant deviations from the expected. Still we do not see a good alternative to using statistical methods for this aspect of model testing.

Dealing with Observational Bias

In early predictive modeling studies, a considerable effort was made to identify the proper statistical methods for model testing (Rose and Altschul 1988; Kvamme 1988, 1990). Not surprisingly, it was concluded that probabilistic sampling is a necessary precondition for collecting data sets to be used for testing as this is the only way to obtain statistically reliable predictions and thus avoid the problem of observational bias. This is not only relevant to questions of representativity of the archaeological data set; it is also important for obtaining representative samples of the independent variables used for prediction. However, various authors concluded that this condition could not be met under most circumstances (Brandt *et al.* 1992; Dalla Bona 1994; Deeben *et al.* 1997) because of biases in archaeological survey procedures and therefore a statistically “valid” prediction of site densities would, in most cases, be illusory. In fact, many predictive modelers for this reason even gave up on inductive modeling and stopped applying statistical tests to judge model performance—which obviously does not solve the problem. Unfortunately, this also meant that some new developments in the statistical sciences were not picked up by predictive modelers (and archaeologists in general).

The development of resampling methods in the 1990s forms a quite radical departure from using traditional, so-called parametric statistical methods (see Efron and Tibshirani 1993; Simon 1997; Lunneborg 2000). Resampling was originally developed for statistical inference from small and non-randomly collected data sets. By using sub-samples from the actual available sample, an unlimited number of new data sets with slightly differing characteristics can be simulated and used for obtaining sample means and confidence intervals and for statistical hypothesis testing. Commonly used methods include bootstrap resampling, permutation resampling, and Monte Carlo simulations. Strange as it may seem to the non-initiated, this method actually works very well and has consistently been shown to produce more reliable estimates of model errors and uncertainty for small and non-randomly collected data sets than using standard parametric methods. Apart from that, the concepts of resampling are straightforward and are usually more easily understood by non-statisticians than traditional statistics—which does not mean that the calculations are always simple. For a science, like archaeology, that has to deal with difficulties in obtaining decent-sized representative data sets in many cases, it offers the potential of obtaining statistically more reliable measures of predictive model quality with relatively small and non-optimal data sets (Verhagen 2007b).

Nevertheless, it remains necessary to always consider the potential problems of survey bias for both development and testing of predictive models. Even though the principles and validity of probabilistic sampling techniques have been recognized by

archaeologists for a long time, everyday practice in current, heritage-management-driven survey is encouraging preferential sampling of particular regions and site types. This is partly due to the fact that the predictive models themselves are predominantly used in CRM for designating zones where survey is enforced (see, *e.g.*, van Leusen 2009), which means that supposedly high probability areas will be surveyed more often than low probability zones. Secondly, it is related to the fact that certain site types are decidedly more difficult to detect than others. For economic reasons, most surveys will therefore tend to favor the detection of relatively large and conspicuous archaeological sites, and this practice is in many cases even codified in national or state survey guidelines. More ephemeral phenomena, like burial grounds, off-site features, and lithic scatters, will be under-represented in most surveys (see Zeidler 1995; Verhagen 2005; Verhagen and Borsboom 2009). And thirdly, because most archaeological survey work is development-driven, surveys will be carried out predominantly in areas with a high level of spatial development. Most of these will be located in the vicinity of urbanized areas, whereas places such as nature reserves, agricultural production zones, or forestry stands will tend to be under-represented not only in the archaeological record but also in the samples of independent variables used for prediction. So, there is a real danger in unwittingly using archaeological survey data sets for predictive model testing, and the most important task is therefore to identify potential biases and try to correct for them. Surprisingly enough, there is very little work done on these issues even though the importance of “source criticism” is acknowledged in many regional archaeological studies (see, *e.g.*, Mischka 2008). However, standard methods for survey data “filtering” do not exist even though the statistical basis for it is well developed and can be relatively easily applied to predictive modeling (see, *e.g.*, Verhagen 2007b; Finke *et al.* 2008; van Leusen *et al.* 2009).

Issues Still to be Solved

It would be far too optimistic to think that applying the cognitive predictive modeling methodology will immediately solve all issues relevant to predicting and interpreting the location of archaeological remains. In the following section, we want to discuss briefly some issues that we think need further study to be better integrated in (cognitive) predictive modeling. These are group agency, temporality, and the prediction of specific occurrences.

Group Agency

Many manifestations of material culture are evidence of group rather than individual actions, but the methodology for dealing with the agency of social collectives seems to be underdeveloped. This is also demonstrated by the majority of agent-based modeling studies where simulated individual agents are provided with rules for social behavior, and patterns will emerge from the interactions of individuals over time. For predictive modeling, however, a certain amount of universality of

individual actions is a necessary precondition, and it begs the question whether the appropriate unit of analysis is the individual or rather a larger social group.

Agency theory in itself does not preclude dealing with groups as agents, but we will have to decide on the scale level first before trying to include agency into our models. As was demonstrated in the preceding section, when trying to model potential activity zones around settlements, the individual perspective is clearly important. The amount of effort that people can put into activities like hunting, fishing, or extracting lithic resources is a primary determinant for territory size (see also, *e.g.*, Trifković 2005). However, the perceived benefit of actually going after these resources may very well be based on collective norms and values. Even more, some of these activities, like hunting big game, can only be performed in groups. Similarly, the choice for a new settlement location or a burial place may be based on group decisions rather than individual preferences. And new groups might bring new technologies and perceptions with them. For example, it is only after Roman colonization that we see widespread quarrying in many mountainous regions of Europe, made possible by technological innovations in mining and road building and triggered by a new, socio-economic demand for building in stone. Again it points to the importance of stipulating generalized individual behaviors within a specific socio-environmental context when trying to use agency in predictive modeling.

Temporality

Clearly, there are techniques available to produce dynamical spatial models based on simulations like agent-based modeling; so, why bother with GIS-based, static cognitive models if we can have the real thing, including agency and temporality? Dynamical models (as far as they have been used by archaeologists) are generally set up for exploratory, heuristic purposes rather than to provide accurate predictions (or more aptly retrodictions) of the actual archaeological record. The practitioners of dynamical modeling are usually very reluctant to draw conclusions in a predictive sense from these models. In most cases, dynamical models will be able to give a good idea of the kinds of pattern that will emerge from certain behaviors but not of the exact location or the chronological order in which they will appear. This is due to the fact that these so-called non-linear models are extremely sensitive to initial conditions and the accumulation of small variations (Allen 1997), a fact that is thought to be true for “real life” as well. Any dynamical model that would aim for (reliable) prediction would therefore have to be based on “real” initial conditions and have benchmark data available of intermediate and end conditions as well in order to limit the outcome of the simulations to more or less realistic scenarios. It will not come as a surprise that this may be a real problem when dealing with archaeological data sets.

Some exceptions are found: Lake (2000), for example, set up a dynamical foraging model for the gathering of hazelnuts on the island of Islay (Scotland) in the Mesolithic. He had to conclude that there was a rather poor fit between the known flint scatters and the modeled areas of hazelnut abundance, casting serious doubts on the theoretical assumption underlying the modeling that hazelnut gathering was a primary determinant for site location. We therefore see real potential in these

techniques for predictive modeling, especially since the cognitive predictive modeling approach does not need to be significantly adapted in order to be used in a dynamical modeling context. For the moment however the technical difficulties involved are still substantial since no off-the-shelf GIS offers dynamical modeling capabilities, and the available open source products have not yet reached a sufficient state of sophistication for easy application. The MAGICAL software developed by Lake for example was developed for use with GRASS version 4 and has not been upgraded since. The SWARM initiative (www.swarm.org), while better maintained, is not coupled to GIS and needs serious programming skills to be used to effect. Perhaps even more telling, on the web pages of the freeware and relatively user friendly agent-based modeling package NetLogo (which has recently acquired the capability to import GIS data), not a single archaeological sample model is mentioned (<http://ccl.northwestern.edu/netlogo/models>). This would point to an only lukewarm interest in these modeling techniques from the side of archaeologists despite the optimism expressed in Bentley and Maschner (2003) and Beekman and Baden (2005)

Predicting the Unusual

Archaeology has sometimes been accused of being obsessed with the unique (Pollard 2004). This evidently has its roots in the old tradition of archaeology as a scientifically justified form of treasure hunting. And even nowadays, few archaeologists will be able to resist the temptation to focus on the more spectacular finds in an excavation and on the most peculiar archaeological sites in a study region. The emphasis on the individual is also evident in current interest in phenomenology and agency even though these approaches deal less directly with the material record. It is one of the most recurrent criticisms of predictive modeling that the models produced will fail to predict unique occurrences—the implication being that these are more interesting and valuable than general occurrences. Note however that Shanks and Tilley (1987, p. 38) questioned the reverse assumption, promoted by New Archaeologists, that the formulation of universal or covering laws is a superior kind of scientific activity.

The question of predicting anomalies however remains a tough nut to crack. Most famously, this problem was identified by Flannery (1976) as the possibility that Teotihuacán might be missed in a probabilistic survey of the Valley of Mexico. While in the case of this particular site there is not much chance that this will happen (“you couldn’t miss it if you tried”), the prediction of anomalies through any form of statistical extrapolation is essentially impossible if we do not transfer hypotheses about their occurrence in other contexts to the area we are interested in. Traditional correlative modeling is no good for this, and even deductive models do not seem to be very well suited for it because of the highly specific characteristics of these sites. Essentially, this also holds true for “negative” anomalies, the places where no archaeological sites are found despite the presence of all conditions favorable to settlement.

However, we want to amend to this that occurrences that are not very common can be predicted even on the basis of a single observation—although the error rate of the prediction will be impossible to establish in such a case. Small numbers always pose

problems when using statistical methods, but it should be mentioned that archaeologists tend to overestimate the number of observations needed for statistical inference by assuming that the ultimate goal of predictive methods is to come up with “robust”, stable estimates with small error ranges. While this is certainly desirable from a management perspective, modern statistical simulation techniques like resampling allow for a realistic estimate of predictive error on the basis of much smaller numbers than are usually considered when applying traditional statistical methods. Ultimately, however, if one builds a model based on an explanatory understanding of cause and effect, then the potential to predict unidentified site locations is directly related to the confidence in the explanation and not the size of the site population.

The power of predictive models to predict the unusual ultimately depends on the data sets used, the spatial extent and resolution chosen, and the theoretical framework applied. For example, Late Bronze Age weapon depositions in fluvial contexts in the Netherlands are typical examples of highly specific sites that are considered extremely valuable for better understanding ritual practices (Roymans and Kortlang 1999; Fontijn 2003). The finds of weaponry (especially swords) in rivers suggest the existence of a warrior elite that may have controlled access to exchange networks with other groups. For some reasons, these weapons were not deposited in graves until later in the Iron Age. They are never found outside fluvial contexts, and given the scarcity of these remains it is assumed that these locations may have had regional or even supra-regional significance as places where competition between local groups could be “played out.” Deposition in rivers is therefore often interpreted by archaeologists as a means to win prestige by destroying valuable objects without forming a threat to the strong collective basis of society as inferred from the burial practices. A second possible interpretation is that these deposition sites were burial places for the elite.

An extrapolation of our current knowledge concerning these depositions would only produce a highly generalized prediction because we cannot get to the specific spatial characteristics of deposition practice itself. Our predictive model may then be accurate in identifying areas (river beds and adjacent flood plains) that are similar to those where depositions have been found in the past, but it is not very precise in delineating potential deposition locations as we do not have very clear ideas about what made people select specific places within flood plains for weapon deposition nor do we have sufficient data to correlate these places to other specific elements of the landscape. Consequently, we will then end up with predictive models with a large amount of “false positives,” *i.e.*, with a low gain.

The Issue of Scale

Other fundamental questions concerning the specificity of predictions are coupled to geographical and temporal scale or resolution. Moving back and forth between scales is a fundamental aspect of archaeological research as all of our generalizations are made on the basis of individual observations that we try to connect into a coherent interpretative and theoretical framework. Much of what we rubricate under problems of scale constitute an artifact of the way in which we collect and classify our data. This is not typical only of archaeology: all cartographic data suffer from

this as they are the end product of generalization and classification schemes. However, it is not always appreciated that the observational scale used may not be the scale at which processes operate and patterns emerge. The modifiable area unit problem for example is a well-known phenomenon in geography: patterns can be made to emerge and disappear by changing the size and positions of the observational units. Furthermore, geographers have long recognized the problem of what is known as ecological fallacy, *i.e.*, the danger of erroneous extrapolation from one scale level to another (see, *e.g.*, Harris 2006).

By choosing a specific scale of analysis for predictive modeling (both in terms of resolution as well as in geographic extent), we already impose limitations on the patterns that may emerge without knowing beforehand what these may be. These are then further limited by our selection of predictor variables. It also means that we can never be certain if we have chosen the “right” scale for our modeling unless we have experimented with varying spatial extents and resolutions. This is a reality that we have to live with and makes it difficult to specify predictive model quality requirements beforehand. In most cases, we will have very little to choose from because of the limitations of the available data sets. One of the reasons why working with digital elevation models and derived variables like slope and aspect became so popular in North American predictive modeling was the availability of cheap or even free elevation models from the US government. Similarly, the spatial extent of predictive models is often limited to administrative boundaries, which has the added disadvantage of introducing edge effects when using variables based on distance calculations.

Conclusions

We hope to have shown that, by considering archaeological predictive modeling from a theoretically informed perspective, both CRM and academic research can benefit. The (deductive) cognitive modeling framework is extremely flexible, easier to operate and understand, better suited for testing purposes, and as far as we can tell produces better predictions than the currently prevailing alternatives—although we have to be a bit wary about the last statement as controlled field tests of predictive models are still very rare. We still see an important role however for inductive, correlative methods as tools for exploratory data analysis in the phase leading up to theory building—after all, theories do not fall out of thin air; they are always based on particular observations. Inductive methods can also be useful for filling in the gaps when theory is un(der)developed. In cases where we insufficiently understand the reasons for a particular behavior to define cause-and-effect relationships, statistical methods can at least give us an idea of the bandwidths involved.

Explanatory models like the Georgia Coast Model not only increase the power of the predictions made but can also be used to explore all kinds of questions on the spatial behavior of people in the past. Predictive modeling in this way evolves into a heuristic tool for developing and testing archaeological theory, not just an equation that will give us an estimate of site densities for a particular region. The “stories” that will emerge from the modeling will then become part of the interpretative stage of archaeological research and may even contribute to, for example, the development of a cultural biography of a particular region for CRM purposes. However, we also

hope to have made clear that the translation of archaeological theory into a predictive model is not always easy and will in most cases need a thorough analysis and re-interpretation of the available theoretical perspectives on the spatial patterning of archaeological remains into practical, middle range theoretical models.

We also think that the methodology suggested in this paper can be helpful in breaking through the self-reinforcing mechanism of site-based archaeological research in CRM that leads to doing more of the same all of the time. Cognitive predictive modeling is flexible enough to accommodate the prediction of all kinds of human activity leading to a variety of material imprints in the archaeological record. Of course, the chain of cause and effect can be carried further to include our evaluating of archaeological sites for their significance to heritage management, our reactions with different kinds of preservation or management behavior, and the results of our efforts on understanding specific cultural landscapes and the archaeological record in general. However, we still have to see whether such an approach may lead to a change in dealing with predictive modeling in CRM. After all, many of the users of predictive maps working in urban and rural planning are not really interested in ever more complex and subtle mapping but need clear guidelines on what to do with the archaeological “problem”.

This may not always be the case however. As part of their land management responsibilities under Section 110 of the National Historic Preservation Act, the Vicksburg District of the US Army Corps of Engineers (USACE) is funding the implementation of a predictive model built on the framework of the Georgia Coast Model. This includes developing a series of interpretive surfaces based on paleoeconomic modeling and aimed at specific temporal periods, site types, and behaviors. Although the study will cover the nearly 43,000,000 acres of the district, the goal is not to simplify the results into a high–low dichotomous model but to provide a quantitative measure for the probability of encountering the classes of sites and behaviors that the regulatory agencies (*i.e.*, the USACE and the Mississippi, Louisiana, and Arkansas State Historic Preservation Offices) have helped to define. This may entail thinking about CRM predictive modeling in a new way—as a tool for both planning future impacts to unsurveyed areas and managing the resources once they are identified. It also entails working in a partnership with agencies and individuals that may have different goals in mind and finding ways in which to bridge the gap between interpretation and useful application.

Acknowledgments The writing of this paper was partly made possible by a VENI research grant from NWO (the Netherlands Organization for Scientific Research) awarded to the first author. We also want to thank Dr. Maaik Groot and Prof. Nico Roymans (both at VU University, Amsterdam) for reading and commenting on an earlier draft of this paper. The paper was further improved by the comments of three anonymous reviewers. All errors of course remain our own.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aldenderfer, M., & Maschner, H. D. G. (Eds.). (1996). *Anthropology, space and geographic information systems*. New York: Oxford University Press.
- Allen, P. M. (1997). Towards a new science of history. In S. E. van der Leeuw & J. McGlade (Eds.), *Time, process and structured transformation in archaeology* (pp. 39–56). London: Routledge.

- Allen, K. M., Green, S. W., & Zubrow, E. B. W. (Eds.). (1990). *Interpreting space: GIS and archaeology*. London: Taylor and Francis.
- Altschul, J. H., Sebastian, L., & Heidelberg, K. (2004). *Predictive modeling in the military. Similar goals, divergent paths*. Rio Rancho: SRI Foundation.
- Axtell, R. L., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlund, A. C., Harbuzer, J., et al. (2006). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. In J. M. Epstein (Ed.), *Generative social science* (pp. 117–129). Princeton: Princeton University Press.
- Barker, G. (1985). *Prehistoric farming in Europe*. Cambridge: Cambridge University Press.
- Barrett, J. C. (2000). A thesis on agency. In M.-A. Dobres & J. E. Robb (Eds.), *Agency in archaeology* (pp. 61–68). London: Routledge.
- Beekman, C. S., & Baden, W. W. (Eds.). (2005). *Nonlinear models for archaeology and anthropology: Continuing the revolution*. Aldershot: Ashgate.
- Bell, J. A. (1994). *Reconstructing prehistory: Scientific method in archaeology*. Philadelphia: Temple University Press.
- Bentley, R. A., & Maschner, H. D. G. (Eds.). (2003). *Complex systems and archaeology. Empirical and theoretical applications*. Salt Lake: University of Utah.
- Besnard, P., & Hanks, S. (1995). *Uncertainty in artificial intelligence* (Vol. 11). San Mateo: Morgan Kaufmann.
- Bettinger, R. L. (1980). Explanatory/predictive models of hunter-gatherer adaptation. In M. B. Schiffer (Ed.), *Advances in archaeological method and theory* (Vol. 3, pp. 155–189). New York: Academic.
- Bettinger, R. L., Malhi, R., & McCarthy, H. (1997). Central place models of acorn and mussel processing. *Journal of Archaeological Science*, 24, 887–899.
- Binford, L. R. (1972). *An archaeological perspective*. New York: Seminar Press.
- Binford, L. R. (1977). General introduction. In L. R. Binford (Ed.), *For theory building in archaeology, essays on faunal remains, aquatic resources, spatial analysis, and systemic modeling* (pp. 1–10). New York: Academic.
- Binford, L. R. (1981). *Bones: Ancient men and modern myths*. New York: Academic.
- Binford, L. R. (1982). Objectivity—explanation—archaeology 1981. In C. Renfrew, M. J. Rowlands, & B. A. Segraves (Eds.), *Theory and explanation in archaeology* (pp. 125–138). New York: Academic.
- Bintliff, J. (1999). Settlement and territory. In G. Barker (Ed.), *The Routledge companion encyclopedia of archaeology* (pp. 505–545). London: Routledge.
- Bird, D. W. (1997). Behavioral ecology and the archaeological consequences of central place foraging among the Merriam. In C. M. Barton & G. A. Clark (Eds.), *Rediscovering Darwin: Evolutionary theory and archaeological explanation* (pp. 291–306). Arlington: American Anthropological Association.
- Bogaard, A. (2004). *Neolithic farming in Central Europe. An archaeobotanical study of crop husbandry practices*. London: Routledge.
- Bonham-Carter, G. F. (1994). *Geographic information systems for geoscientists. Modelling with GIS*. Oxford: Pergamon Press.
- Boserup, E. (1965). *The conditions of agricultural growth*. London: Allen and Unwin.
- Brandt, R. W., Groenewoudt, B. J., & Kvamme, K. L. (1992). An experiment in archaeological site location: Modelling in the Netherlands using GIS techniques. *World Archaeology*, 24, 268–282.
- Burns, G., Fronabarger, A. K., & Whitley, T. G. (2008). Predictive modeling of cultural resources in the Theban Necropolis, Luxor, Egypt. In A. Posluschny, K. Lambers, & I. Herzog (Eds.), *Layers of perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007* (pp. 307 & CD-ROM). Bonn: Dr. Rudolf Habelt GmbH.
- Carlstein, T. (1982). *Time resources, society and ecology: On the capacity for human interaction in space and time. Vol. 1: Preindustrial societies*. London: Allen & Unwin.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon.
- Christaller, W. (1935). *Die Zentralen Orte in Süddeutschland*. Jena: J. Fischer Verlag.
- Cleuziou, S. (2007). Evolution toward complexity in a coastal desert environment: The early bronze age in the Ja'alan, Sultanate of Oman. In T. A. Kohler & S. E. van der Leeuw (Eds.), *Model-based archaeology of socionatural systems* (pp. 213–231). Santa Fe: School of Advanced Research.
- Cohen, R. (Ed.). (1985). *The development of spatial cognition*. Hillsdale: Erlbaum Lawrence.
- Coolbaugh, M. F., Raines, G. L., & Zehner, R. E. (2007). Assessment of exploration bias in data-driven predictive models and the estimation of undiscovered resources. *Natural Resources Research*, 16, 199–207.

- Cowgill, G. L. (2000). "Rationality" and contexts in agency theory. In M.-A. Dobres & J. E. Robb (Eds.), *Agency in archaeology* (pp. 51–60). London: Routledge.
- Cox, D. R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society, Series A*, 155, 291–301.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., et al. (2007). Random forests for classification in ecology. *Ecology*, 88, 2783–2792.
- Dalla Bona, L. (1994). *Ontario Ministry of Natural Resources Archaeological Predictive Modelling Project*. Thunder Bay: Center for Archaeological Resource Prediction, Lakehead University.
- de Hingh AE (2000) Food production and food procurement in the bronze age and early iron age (2000–500 BC). The organisation of a diversified and intensified agrarian system in the Meuse–Demerscheldt region (The Netherlands and Belgium) and the region of the river Moselle (Luxemburg and France). Leiden: Faculty of Archaeology, Leiden University
- Dean, J. S., Gumerman, G. J., Epstein, J. M., Axtell, R. L., Swedlund, A. C., Parker, M. T., et al. (2006). Understanding Anasazi culture change through agent-based modeling. In J. M. Epstein (Ed.), *Generative social science* (pp. 90–116). Princeton: Princeton University Press.
- Deeben, J., Hallewas, D., Kolen, J., & Wiemer, R. (1997). Beyond the crystal ball: Predictive modelling as a tool in archaeological heritage management and occupation history. In W. Willems, H. Kars, & D. Hallewas (Eds.), *Archaeological heritage management in the Netherlands. Fifty years state service for archaeological investigations* (pp. 76–118). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Deeben, J., Hallewas, D. P., & Maarleveld, T. J. (2002). Predictive modelling in archaeological heritage management of the Netherlands: The indicative map of archaeological values (2nd generation). *Berichten van de Rijksdienst voor het Oudheidkundig Bodemonderzoek*, 45, 9–56.
- Dobres, M.-A., & Robb, J. E. (2000). Agency in archaeology: Paradigm or platitude? In M.-A. Dobres & J. E. Robb (Eds.), *Agency in archaeology* (pp. 3–17). London: Routledge.
- Dobres, M.-A., & Robb, J. E. (2005). Doing "Agency": Introductory remarks on methodology. *Journal of Archaeological Method and Theory*, 12, 159–166.
- Dorman, J. L. (2002). Agency and archaeology: Past, present and future directions. *Journal of Archaeological Method and Theory*, 9, 303–329.
- Downs, R. M., & Stea, D. (Eds.). (1973). *Image and environment*. Chicago: Aldine.
- Downs, R. M., & Stea, D. (1977). *Maps in minds: Reflections on cognitive mapping*. New York: Harper Row.
- Ducke, B., & Münch, U. (2005). Predictive modelling and the archaeological heritage of Brandenburg (Germany). In M. van Leusen & H. Kamermans (Eds.), *Predictive modelling for archaeological heritage management: A research agenda* (pp. 93–108). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Ejstrud, B. (2003). Indicative models in landscape management: Testing the methods. In J. Kunow & J. Müller (Eds.), *Symposium on the archaeology of landscapes and geographic information systems. Predictive maps, settlement dynamics and space and territory in prehistory* (pp. 119–134). Wünsdorf: Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Landesmuseum.
- Ejstrud, B. (2005). Taphomic models: Using Dempster–Shafer theory to assess the quality of archaeological data and indicative models. In M. van Leusen & H. Kamermans (Eds.), *Predictive modelling for archaeological heritage management: A research agenda* (pp. 183–194). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Ellen, R. (1982). *Environment, subsistence and system. The ecology of small-scale social systems*. Cambridge: Cambridge University Press.
- Emlen, J. M. (1966). The role of time and energy in food preference. *The American Naturalist*, 100, 611–617.
- Epstein, J. (Ed.). (2006). *Generative social science*. Princeton: Princeton University Press.
- Ericson, J. E., & Goldstein, R. (1980). Work space: A new approach to the analysis of energy expenditure within site catchments. In F. J. Findlow & J. E. Ericson (Eds.), *Catchment analysis. Essays on prehistoric resource space* (pp. 21–30). Los Angeles: Department of Anthropology, University of California.
- Favory, F., & van der Leeuw, S. (1998). ARCHAEOMEDES, la dynamique spatio-temporelle de l'habitat antique dans la Vallée du Rhône: bilan et perspectives. *Revue Archéologique de Narbonnaise*, 31, 257–298.
- Finke, P. A., Meylemans, E., & Van De Wauw, J. (2008). Mapping the possible occurrence of archaeological sites by Bayesian inference. *Journal of Archaeological Science*, 35, 2786–2796.
- Flannery, K. V. (1976). *The early Meso-American village*. New York: Academic.

- Fleming, A. (2006). Post-processual landscape archaeology: A critique. *Cambridge Archaeological Journal*, 16, 267–280.
- Fokkens, H. (1998). *Drowned landscape. The occupation of the western part of the Frisian–Drentian plateau, 4400 BC–AD 500*. Assen: Van Gorcum.
- Fokkens, H. (2002). Vee en voorouders: Centrale elementen uit het dagelijks leven in de bronstijd. In H. Fokkens & R. Jansen (Eds.), *2000 Jaar bewoningsdynamiek. Brons- en ijzertijdbewoning in het Maas-Demer-Scheldegebied* (pp. 126–148). Leiden: Faculteit der Archeologie, Universiteit Leiden.
- Fontijn, D. (2003). *Sacrificial landscapes cultural biographies of persons, objects and 'natural' places in the Bronze Age of the Southern Netherlands, c. 2300–600 BC*. Leiden: Leiden University Press.
- Gärting, T., & Evans, G. W. (Eds.). (1991). *Environment, cognition, and action—an integrated approach*. New York: Oxford University Press.
- Georgia Department of Natural Resources (DNR) (2010). Wildlife Resources Division: Educator resources; species fact sheets. <http://www.georgiawildlife.org/education/educator-resources>. Accessed 2 November 2010.
- Gerritsen, F. (2003). *Local identities. Landscape and community in the late prehistoric Meuse–Demer–Scheldt region*. Amsterdam: Amsterdam University Press.
- Gimblett, H. R. (Ed.). (2002). *Integrating geographic information systems and agent-based modeling techniques*. Oxford: Oxford University Press.
- Goodchild, H. (2006). Modelling agricultural production. A methodology for predicting land use and populations. In J. F. Haldon (Ed.), *General issues in the study of Medieval logistics* (pp. 199–228). Leiden: Brill.
- Grayson, D. K., & Cannon, M. D. (1999). Human paleoecology and foraging theory in the great basin. In C. Beck (Ed.), *Models for the millennium* (pp. 141–151). Salt Lake: University of Utah Press.
- Grayson, D. K., & Delpech, J. (1998). Changing diet breadth in the early upper paleolithic of Southwestern France. *Journal of Archaeological Science*, 25, 1119–1130.
- Green, E. L. (1973). Location analysis of prehistoric Maya sites in Northern British Honduras. *American Antiquity*, 38(3), 279–293.
- Gregg, S. A. (1988). *Foragers and farmers. Population interaction and agricultural expansion in Prehistoric Europe*. Chicago: The University of Chicago.
- Gumerman, G. J., Swedlund, A. C., Dean, J. S., & Epstein, J. M. (2006). The evolution of social behavior in the prehistoric American Southwest. In J. M. Epstein (Ed.), *Generative social science* (pp. 130–143). Princeton: Princeton University Press.
- Hames, R., & Vickers, W. (1982). Optimal diet breadth theory as a model to explain variability in Amazonian hunting. *American Ethnologist*, 9, 358–378.
- Harris, T. M. (2006). Scale as artifact: GIS, ecological fallacy, and archaeological analysis. In G. Lock & B. L. Molyneux (Eds.), *Confronting scale in archaeology. Issues of theory and practice* (pp. 39–53). New York: Springer.
- Harris, T., & Lock, G. (1995). Toward an evaluation of GIS in European archaeology: The past, present and future of theory and applications. In G. Lock & Z. Staničić (Eds.), *Archaeology and geographic information systems: A European perspective* (pp. 349–366). London: Taylor and Francis.
- Hegmon, M. (2003). Setting theoretical egos aside: Issues and theory in North American archaeology. *American Antiquity*, 68(2), 213–243.
- Hempel, C. G. (1965). *Aspects of scientific explanation*. New York: Free Press.
- Hengl, T., Sierdema, H., Radović, A., & Dilod, A. (2009). Spatial prediction of species' distributions from occurrence-only records: Combining point pattern analysis, ENFA and regression–kriging. *Ecological Modelling*, 220(24), 3499–3511.
- Hitchcock, C. R. (1996). Causal decision theory and decision theoretic causation. *Nous*, 30, 508–526.
- Hodder, I. (1991). *Reading the past: Current approaches to interpretation in archaeology* (2nd ed.). Cambridge: Cambridge University Press.
- Hoving, C. L., Harrison, D. J., Krohn, W. B., Jakubas, W. J., & McCullough, M. A. (2004). Canada lynx *Lynx canadensis* habitat and forest succession in northern Maine, USA. *Wildlife Biology*, 10(4), 285–294.
- Hudak, G. J., E. Hobbs, A. Brooks, C. A. Sersland, C. Phillips (eds.). 2002. Mn/model final report 2002: A predictive model of precontact archaeological site location for the state of Minnesota. St. Paul: Minnesota Department of Transportation. http://www.mnmodel.dot.state.mn.us/pages/final_report.html. Accessed on 2 November 2010.
- Hume, D. (1739). *A treatise of human nature*. London: John Noon.
- Huston, M. A. (2002). Introductory essay: Critical issues for improving predictions. In J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, & F. B. Samson (Eds.), *Predicting species occurrences: Issues of accuracy and scale* (pp. 7–21). Washington DC: Island.

- Ingold, T. (1993). The temporality of the landscape. *World Archaeology*, 25, 152–174.
- Ingold, T. (2000). *The perception of the environment. Essays on livelihood, dwelling and skill*. London: Routledge.
- Ittleson, W. H. (Ed.). (1973). *Environment and cognition*. New York: Seminar.
- Jochim, M. (1976). *Hunter-gatherer subsistence and settlement: A predictive model*. New York: Academic.
- Jochim, M. (1981). *Strategies for survival: Cultural behavior in an ecological context*. New York: Academic.
- Johnson, M. H. (2006). On the nature of theoretical archaeology and archaeological theory. *Archaeological Dialogues*, 13(2), 117–132.
- Jones, K. T., & Madsen, D. B. (1989). Calculating the cost of resource transportation. *Current Anthropology*, 30, 529–534.
- Judge, J. W., & Sebastian, L. (Eds.). (1988). *Quantifying the present and predicting the past: Theory, method and application of archaeological predictive modelling*. Denver: U.S. Department of the Interior, Bureau of Land Management.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, XLVII, 263–291.
- Kamermand, H., & Wansleben, M. (1999). Predictive modelling in Dutch archaeology, joining forces. In J. A. Barceló, I. Briz, & A. Vila (Eds.), *New techniques for old times—CAA98. Computer applications and quantitative methods in archaeology* (pp. 225–230). Oxford: Archaeopress.
- Kamermand, H., van Leusen, M., & Verhagen, P. (Eds.). (2009). *Archaeological prediction and risk management*. Leiden: Leiden University Press.
- Kelly, R. L. (1995). *The foraging spectrum: Diversity in hunter-gatherer lifeways*. Washington DC: Smithsonian Institution.
- King, T. F. (1984). An overview of the archaeological law at the federal level. *American Archaeology*, 4, 115–118.
- Kitchin, R., & Blades, M. (2002). *The cognition of geographic space*. London: I.B. Taurus.
- Kitchin, R., & Freundschuh, S. (Eds.). (2000). *Cognitive mapping. Past, present, and future*. London: Routledge.
- Kohler, T. A., & Gumerman, G. J. (Eds.). (2000). *Dynamics in human and primate societies: Agent-based modeling of social and spatial processes*. New York: Oxford University Press.
- Kohler, T. A., & Parker, S. C. (1986). Predictive models for archaeological resource location. In M. B. Schiffer (Ed.), *Advances in archaeological method and theory* (Vol. 9, pp. 397–452). New York: Academic.
- Kohler, T. A., & van der Leeuw, S. (Eds.). (2007). *Model-based archaeology of socio-natural systems*. Santa Fe: School of Advanced Research.
- Kohler, T. A., Johnson, C. D., Varien, M., Ortman, S., Reynolds, R., Kobti, Z., et al. (2007). Settlement ecodynamics in the prehispanic Central Mesa Verde Region. In T. A. Kohler & S. E. van der Leeuw (Eds.), *Model-based archaeology of socio-natural systems* (pp. 61–104). Santa Fe: School of Advanced Research.
- Kosso, P. (1991). Method in archaeology: Middle-range theory as hermeneutics. *American Antiquity*, 56, 621–627.
- Kosso, P. (2006). Introduction. The epistemology of archaeology. In G. G. Fagan (Ed.), *Archaeological fantasies* (pp. 3–22). London: Routledge.
- Kvamme, K. L. (1983). *A manual for predictive site location models: examples from the Grand Junction District, Colorado*. Grand Junction District: Bureau of Land Management.
- Kvamme, K. L. (1984). Models of prehistoric site location near Pinyon Canyon, Colorado. In C. J. Condie (Ed.), *Papers of the Philmont Conference on the archaeology of Northeastern New Mexico* (pp. 349–370). Albuquerque: New Mexico Archaeological Council.
- Kvamme, K. L. (1985). Determining empirical relationships between the natural environment and prehistoric site locations: A hunter-gatherer example. In C. Carr (Ed.), *For concordance in archaeological analysis: Bridging data structure, quantitative technique, and theory* (pp. 208–238). Prospect Heights: Waveland.
- Kvamme, K. L. (1988). Development and testing of quantitative models. In W. J. Judge & 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. L. (1990). The fundamental principles and practice of predictive archaeological modeling. In A. Voorrips (Ed.), *Mathematics and information science in archaeology: A flexible framework* (pp. 257–295). Bonn: Holos Verlag.

- Lake, M. W. (2000). MAGICAL computer simulation of mesolithic foraging. In T. A. Kohler & G. J. Gumerman (Eds.), *Dynamics in human and primate societies: Agent-based modeling of social and spatial processes* (pp. 107–143). New York: Oxford University Press.
- Levinson, L. C. (1992). *Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr*. Nijmegen: The Max Planck Institute for Psycholinguistics.
- Limp, W. F., & Carr, C. (1985). The analysis of decision making: Alternative applications in archaeology. In C. Carr (Ed.), *For concordance in archaeological analysis: Bridging data structure, quantitative technique, and theory* (pp. 128–172). Prospect Heights: Waveland.
- Llobera, M. (1996). Exploring the topography of mind: GIS, social space and archaeology. *Antiquity*, 70, 612–622.
- Llobera, M. (2000). Understanding movement: A pilot model towards the sociology of movement. In G. Lock (Ed.), *Beyond the map. Archaeology and spatial technologies* (pp. 65–84). Amsterdam: IOS Press/Ohmsha.
- Llobera, M. (2001). Building past landscape perception with GIS: Understanding topographic prominence. *Journal of Archaeological Science*, 28, 1005–1014.
- Llobera, M. (2003). Extending GIS-based visual analysis: The concept of visualsapes. *International Journal of Geographical Information Science*, 17, 25–48.
- Lock, G., & Stančić, Z. (Eds.). (1995). *Archaeology and geographic information systems: A European perspective*. London: Taylor and Francis.
- Lunneborg, C. E. (2000). *Data analysis by resampling: Concepts and applications*. Pacific Grove: Duxbury.
- MacArthur, R. H., & Pianka, E. R. (1966). On the optimal use of a patchy environment. *The American Naturalist*, 100, 603–609.
- Mathys, L., Zimmermann, N. E., Zbinden, N., & Suter, W. (2006). Identifying habitat suitability for hazel grouse *Bonasia bonasia* at the landscape scale. *Wildlife Biology*, 12(4), 357–366.
- Merton, R. K. (1968). *Social theory and social structure* (3rd ed.). New York: Free.
- Metcalfe, D., & Barlow, K. R. (1992). A model for exploring the optimal trade-off between field processing and transport. *American Anthropologist*, 94, 340–356.
- Millard, A. R. (2005). What can Bayesian statistics do for predictive modelling? In M. van Leusen & H. Kamermans (Eds.), *Predictive modelling for archaeological heritage management: A research agenda* (pp. 169–182). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton: Princeton University Press.
- Mischka, D. (2008). Territorial modelling and archaeological data: How complete must the picture be? In A. Posluschny, K. Lambers, & I. Herzog (Eds.), *Layers of perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007* (pp. 280 & CD-ROM). Bonn: Dr. Rudolf Habelt GmbH.
- Moore, G. T. (1979). Knowing about environmental knowing: The current state of theory and research on environmental cognition. *Environment and Behavior*, 11, 33–70.
- Nackaerts, K., Govers, G., & Orshoven, J. V. (1999). Accuracy assessment of probabilistic visibilities. *International Journal of Geographical Information Science*, 13, 709–721.
- Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation*. New York: Harcourt, Brace and World.
- Natural Resources Spatial Analysis Lab (NARSAL), University of Georgia (2010). Georgia Gap Analysis Project (GA_GAP): Vertebrate Models. http://narsal.uga.edu/gap/vertebrate_models. Website with restricted access.
- O'Connell, J. F., & Hawkes, K. (1984). Food choice and foraging sites among the Alyawara. *Journal of Anthropological Research*, 40, 435–504.
- Oliver, R. M., & Smith, J. Q. (1990). *Influence diagrams, belief nets, and decision analysis*. New York: Wiley.
- Oreskes, N. (1998). Evaluation (not validation) of quantitative models. *Environmental Health Perspectives*, 106, 1453–1460.
- Oreskes, N. (2003). The role of quantitative models in science. In C. D. Canham, J. J. Cole, & W. K. Lauenroth (Eds.), *Models in ecosystem science* (pp. 13–31). Princeton: Princeton University Press.
- Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641–646.
- Orians, G. F., & Pearson, N. E. (1979). On the theory of central place foraging. In D. J. Horn, R. D. Mitchell, & C. R. Stairs (Eds.), *Analysis of ecological systems* (pp. 154–177). Columbus: Ohio State University Press.

- O'Sullivan, D., and M.N. Gahegan. 2007. Issues around verification, validation, calibration, and confirmation of agent-based models of complex spatial systems. Paper presented at the NSF-ESRC Workshop on agent-based models of complex spatial systems. Santa Barbara, CA, April 14–16. <http://www.m8p.com>. Accessed on 2 November 2010.
- Parker, S. C. (1985). Predictive modeling of site settlement systems using multivariate logistics. In C. Carr (Ed.), *For concordance in archaeological analysis: Bridging data structure, quantitative technique, and theory* (pp. 173–207). Prospect Heights: Waveland Press.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge: Cambridge University Press.
- Peeters, J. H. M. (2007). *Hoge Vaart-A27 in context: Towards a model of Mesolithic–Neolithic land use dynamics as a framework for archaeological heritage management*. Amersfoort: Rijksdienst voor Archeologie, cultuurlandschap en monumenten.
- Peters, J., Baets, B. De, Verhoest, N. E. C., Samson, R., Degroev, S., Becker, P. De, et al. (2007). Random forests as a tool for ecohydrological distribution modelling. *Ecological Modelling*, 207, 304–318.
- Plog, F., & Hill, J. M. (1971). Explaining variability in the distribution of sites. In G. J. Gumerman (Ed.), *The distribution of prehistoric population aggregates* (pp. 7–37). Prescott: Prescott College Press.
- Pollard, A. M. (2004). Putting infinity up on trial: A consideration of the role of scientific thinking in future archaeologies. In J. Bintliff (Ed.), *A companion to archaeology* (pp. 380–396). Malden: Blackwell Publishing Ltd.
- Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
- Raab, L. M., & Goodyear, A. C. (1984). Middle-range theory in archaeology: A critical review of origins and applications. *American Antiquity*, 49, 255–268.
- Raines, G. L. (1999). Evaluation of weights of evidence to predict epithermal-gold deposits in the great basin of the Western United States. *Natural Resources Research*, 8, 257–276.
- Reitz, E. J., Pavao-Zuckerman, B., Weinand, D. C., & Duncan, G. A. (2010). *Mission and Pueblo of Santa Catalina de Guale, St. Catherines Island, Georgia: A comparative zooarchaeological analysis*. New York: American Museum of Natural History.
- Renfrew, C., & Zubrow, E. B. (Eds.). (1994). *The ancient mind. Elements of cognitive archaeology*. Cambridge: Cambridge University Press.
- Robb, J., & Van Hove, D. (2003). Gardening, foraging and herding: Neolithic land use and social territories in Southern Italy. *Antiquity*, 77, 241–254.
- Rose, M. R., & Altschul, J. H. (1988). An overview of statistical method and theory for quantitative model building. In W. J. Judge & L. Sebastian (Eds.), *Quantifying the present and predicting the past: Theory, method, and application of archaeological predictive modelling* (pp. 173–256). Denver: US Department of the Interior, Bureau of Land Management.
- Roymans, N., & Gerritsen, F. (2002). Landscape, ecology and mentalités: A long-term perspective on developments in the Meuse–Demer–Scheldt region. *Proceedings of the Prehistoric Society*, 68, 257–287.
- Roymans, N., & Kortlang, F. (1999). Urnfield symbolism, ancestors and the land in the Lower Rhine region. In F. Theuvs & N. Roymans (Eds.), *Land and ancestors. Cultural dynamics in the Urnfield period and the Middle Ages in the Southern Netherlands* (pp. 33–61). Amsterdam: Amsterdam University Press.
- Roymans, N., & Theuvs, F. (1999). Long-term perspectives on land and people in the Meuse–Demer–Scheldt region (1100 BC–1500 AD). In F. Theuvs & N. Roymans (Eds.), *Land and ancestors. Cultural dynamics in the Urnfield period and the Middle Ages in the Southern Netherlands* (pp. 1–32). Amsterdam: Amsterdam University Press.
- Roymans, N., Gerritsen, F., van der Heijden, C., Bosma, K., & Kolen, J. (2009). Landscape biography as research strategy: The case of the South Netherlands project. *Landscape Research*, 34, 337–359.
- Rygielski, C., Wang, J.-C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, 24(4), 483–502.
- Saarinen, T. F., Seamon, D., & Sell, J. L. (Eds.). (1984). *Environmental perception and behavior: An inventory and prospect*. Chicago: Department of Geography, University of Chicago.
- Salmon, W. C. (1971). *Statistical explanation and statistical relevance*. Pittsburgh: University of Pittsburgh Press.
- Salmon, W. C. (1998). *Causality and explanation*. Oxford: Oxford University Press.
- Schiffer, M. B. (1976). *Behavioral archaeology*. New York: Academic.
- Sebastian, L., & Judge, J. W. (1988). Predicting the past: Correlation, explanation and the use of archaeological models. In W. J. Judge & L. Sebastian (Eds.), *Quantifying the present and predicting the past: Theory, method, and application of archaeological predictive modelling* (pp. 1–18). Denver: US Department of the Interior, Bureau of Land Management.

- Shanks, M., & Tilley, C. (1987). *Re-constructing archaeology*. Cambridge: Cambridge University Press.
- Shanks, D. R., Holyoak, K. J., & Medin, D. L. (Eds.). (1996). *The psychology of learning and motivation*. New York: Academic.
- Simms, S. R. (1987). *Behavioral ecology and hunter-gatherer foraging: An example from the Great Basin*. Oxford: BAR International Series 381.
- Simon, J.L. 1997. Resampling: The new statistics, 2nd edition. <http://www.resample.com/content/text/index.shtml>. Accessed on 2 November 2010.
- Smith, E. A. (1991). *Inujamuit foraging strategies: Evolutionary ecology of an arctic hunting economy*. New York: Aldine.
- Smith, B. D. (Ed.). (1992). *Rivers of change: Essays on early agriculture in Eastern North America*. Washington, DC: Smithsonian Institution.
- Spek, T., Groenman-van Waateringe, W., Kooistra, M., & Bakker, L. (2003). Formation and land-use history of celtic fields in north-west Europe. An interdisciplinary case study at Zeijen—The Netherlands. *European Journal of Archaeology*, 6, 141–169.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton: Princeton University Press.
- Styles, B. N. (1981). *Faunal exploitation and resource selection: Early Late Woodland subsistence in the lower Illinois Valley*. Evanston: Northwestern University Archeological Program.
- Thomas, D. H. (1978). The awful truth about statistics in archaeology. *American Antiquity*, 43, 231–244.
- Thomas, D. H. (Ed.). (2008). *Native American landscapes of St. Catherines Island, Georgia. Volume I: Theoretical framework*. New York: American Museum of Natural History.
- Thurston, T. L. (2007). Infields, outfields, and broken lands: Agricultural intensification and the ordering of space during Danish state formation. In T. L. Thurston & C. T. Fisher (Eds.), *Seeking a richer harvest. The archaeology of subsistence intensification, innovation and change* (pp. 155–191). New York: Springer.
- Tilley, C. (2004). *The materiality of stone. Explorations in landscape phenomenology I*. Oxford: Berg.
- Tilley, C., Hamilton, S., & Bender, B. (2000). Art and the re-presentation of the past. *Journal of the Royal Anthropological Institute (N.S.)*, 6, 35–62.
- Tobler, W. (1993). Non-isotropic geographic modeling. In W. Tobler (Ed.), *Three presentations on geographic analysis and modeling*. Santa Barbara: National Center for Geographic Information and Analysis, University of California.
- Trifković, V. (2005). *Construction of space in Early Holocene iron gates*. Oxford: The Institute of Archaeology, University of Oxford.
- Tschauner, H. (1996). Middle-range theory, behavioral archaeology, and postempiricist philosophy of science in archaeology. *Journal of Archaeological Method and Theory*, 3, 1–30.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- van der Leeuw, S. E., & McGlade, J. (Eds.). (1997). *Time, process and structured transformation in archaeology*. London: Routledge.
- Van Hove, D. (2003). *Imagining Calabria. A GIS approach to Neolithic landscapes*. Southampton: University of Southampton.
- van Leusen, M. (2009). Archaeological predictions contested: The role of the Dutch indicative map of archaeological values (IKAW) in local planning procedures. In H. Kamermans, M. van Leusen, & P. Verhagen (Eds.), *Archaeological prediction and risk management* (pp. 39–46). Leiden: Leiden University Press.
- van Leusen, M., & Kamermans, H. (Eds.). (2005). *Predictive modelling for archaeological heritage management: A research agenda*. Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- van Leusen, M., Millard, A. R., & Ducke, B. (2009). Dealing with uncertainty in archaeological prediction. In H. Kamermans, M. van Leusen, & P. Verhagen (Eds.), *Archaeological prediction and risk management* (pp. 123–173). Leiden: Leiden University Press.
- Verhagen, P. (2005). Prospection strategies and archaeological predictive modelling. In M. van Leusen & H. Kamermans (Eds.), *Predictive modelling for archaeological heritage management: A research agenda* (pp. 109–121). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Verhagen, P. (2006). Quantifying the qualified: The use of multicriteria methods and Bayesian statistics for the development of archaeological predictive models. In M. Mehrer & K. Wescott (Eds.), *GIS and archaeological site location modeling* (pp. 191–216). Boca Raton: CRC Press.
- Verhagen, P. (2007a). A condensed history of predictive modelling in archaeology. In P. Verhagen (Ed.), *Case studies in archaeological predictive modelling* (pp. 13–25). Leiden: Leiden University Press.
- Verhagen, P. (2007b). Predictive models put to the test. In P. Verhagen (Ed.), *Case studies in archaeological predictive modelling* (pp. 115–168). Leiden: Leiden University Press.

- Verhagen, P., & Borsboom, A. (2009). The design of effective and efficient trial trenching strategies for discovering archaeological sites. *Journal of Archaeological Science*, 36, 1807–1816.
- Verhoef, P. C. (2003). Understanding the effects of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4), 30–45.
- Wakker, P. P., Timmermans, D. R. M., & Machielse, I. A. (2003). *The effects of statistical information on insurance decisions and risk attitudes*. Amsterdam: Department of Economics, University of Amsterdam.
- Warren, R. E. (1990). Predictive modelling in archaeology: A primer. In K. M. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 90–111). London: Taylor and Francis.
- Weber, A. (1929). *Theory of the location of industries*. Chicago: University of Chicago.
- Wescott, K. L., & Brandon, R. J. (Eds.). (2000). *Practical applications of GIS for archaeologists: A predictive modeling kit*. London: Taylor and Francis.
- Wheatley, D. (1993). Going over old ground: GIS, archaeological theory and the act of perception. In J. Andresen, T. Madsen, & I. Scollar (Eds.), *Computing the past. Computer applications and quantitative methods in archaeology. CAA92* (pp. 133–138). Aarhus: Aarhus University Press.
- Wheatley, D. 2004. Making space for an archaeology of place. *Internet Archaeology* 15. http://intarch.ac.uk/journal/issue15/wheatley_index.html.
- Wheatley, D., & Gillings, M. (2002). *Spatial technology and archaeology. The archaeological applications of GIS*. London: Taylor and Francis.
- Whitley, T. G. (2003). GIS as an interpretative tool for addressing risk management and cognitive spatial dynamics in a slave society. In M. Doerr & A. Sarris (Eds.), *CAA 2002. The digital heritage of archaeology. Computer applications and quantitative methods in archaeology. Proceedings of the 30th Conference, Heraklion, Crete, April 2002, archive of monuments and publications* (pp. 209–215). Heraklion: Hellenic Ministry of Culture.
- Whitley, T. (2004). Causality and cross-purposes in archaeological predictive modeling. In A. Fischer Ausserer, W. Börner, M. Goriani, & L. Karlhuber-Vöckl (Eds.), *[Enter the past]: The E-way into the four dimensions of cultural heritage: CAA 2003: Computer applications and quantitative methods in archaeology: Proceedings of the 31th Conference, Vienna, Austria, April 2003* (pp. 236–239). Oxford: Archaeopress.
- Whitley, T. G. (2005). A brief outline of causality-based cognitive archaeological probabilistic modeling. In M. van Leusen & H. Kamermans (Eds.), *Predictive modelling for archaeological heritage management: A research agenda* (pp. 123–138). Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Whitley, T. G., Moore, G., Goel, G., & Jackson, D. (2009). Beyond the marsh: Settlement choice, perception and spatial decision-making on the Georgia coastal plain. In B. Frischer, J. Crawford, & D. Kollers (Eds.), *CAA 2009. Making history interactive. computer applications and quantitative methods in archaeology. Proceedings of the 37th Conference, Williamsburg, VA, USA* (pp. 380–390). Oxford: Archaeopress.
- Wilkinson, T. J., Gibson, M., Christianson, J. H., Widell, M., Schloen, D., Kouchoukos, N., et al. (2007). Modeling settlement systems in a dynamic environment: Case studies from Mesopotamia. In T. A. Kohler & S. E. van der Leeuw (Eds.), *Model-based archaeology of socio-natural systems* (pp. 175–208). Santa Fe: School of Advanced Research.
- Wiley, G. R. (1953). *Prehistoric settlement in the Virú Valley, Peru*. Washington, DC: United States Government Printing Office.
- Winterhalder, B. (1987). The analysis of hunter–gatherer diets: Stalking and optimal foraging model. In M. Harris & E. Ross (Eds.), *Food and evolution* (pp. 311–319). Philadelphia: Temple University Press.
- Winterhalder, B., & Kennett, D. (2006). Behavioral ecology and the transition from hunting and gathering to agriculture. In D. Kennett & B. Winterhalder (Eds.), *Behavioral ecology and the transition to agriculture* (pp. 1–21). Berkeley: University of California Press.
- Witcher, R. E. (1999). GIS and landscapes of perception. In M. Gillings, D. Mattingly, & J. van Dalen (Eds.), *Geographical information systems and landscape archaeology* (pp. 13–22). Oxford: Oxbow.
- Wylie, A. (2002). *Thinking from things. Essays in the philosophy of archaeology*. Berkeley: University of California Press.
- Zeanah, D. (1996). *Predicting settlement patterns and mobility strategies: An optimal foraging analysis of hunter–gatherer use of mountain, desert, and wetland habitats in the Carson Desert*. Salt Lake: Department of Anthropology, University of Utah.
- Zeidler, J. A. (1995). *Archaeological inventory survey standards and cost-estimation guidelines for the Department of Defense*. Champaign: USACERL.

- Zeidler, J. A. (Ed.). (2001). *Dynamic modeling of landscape evolution and archaeological site distributions: A three-dimensional approach*. Fort Collins: Center for Environmental Management of Military Lands, Colorado State University.
- Zimmermann, F., & Breitenmoser, U. (2007). Potential distribution and population size of the Eurasian lynx *Lynx lynx* in the Jura Mountains and possible corridors to adjacent ranges. *Wildlife Biology*, 13(4), 406–416.