Intensities, interactions and uncertainties: some new approaches to archaeological distributions

Andrew Bevan, Enrico Crema, Xiuzhen Li, Alessio Palmisano
(UCL Institute of Archaeology)

Postprint of a chapter in Bevan, A. and Lake, M. (eds.), *Computational Approaches to Archaeological Spaces*, Walnut Creek: Left Coast Press.

1. Introduction

While distribution maps are nearly as old as the discipline of archaeology itself, most archaeologists still rely on personal intuition with regard to their assessment both of the spatial patterns they recover and the environmental processes and human behaviours that might be behind these patterns. To some extent, this general preference for intuitive readings of space in the archaeological record probably reflects several decades of disillusionment with quantitative spatial methods, after a flurry of early interest during the 1970s (e.g. Hodder and Orton 1976; Clarke 1977), and a continuing wish to prioritise the study of cultural spaces as subjectively experienced and meaningfully constituted by their human inhabitants (e.g. Gregory and Urry eds. 1985). Interestingly, even the enthusiastic uptake of Geographic Information Systems (GIS) from the 1990s onwards did little to change this situation with regard to spatial pattern analysis, as most off-the-shelf GIS software was targeted at data management and querying, digital cartography and enhanced visualisation, as well as certain focused modelling agendas (e.g. terrain, visibility and movement). Effectively, the study of distribution maps in archaeology merely carried on as it was, with a healthy dose of expert intuition, and perhaps in slightly richer visual form.

However, while human involvement in the act of interpretation is undeniably a crucial and enduring aspect of archaeological research, there remain good reasons to characterise spatial distributions in more formal, quantitative ways. This paper focuses on a set of point pattern and process models that, we argue, now puts archaeologists in a position to return to the analysis of spatial pattern and process with renewed ambition, especially with regard to distribution maps. The first section below considers current theoretical approaches to point distributions and subsequent sections then address three cases studies that highlight some important conceptual issues and new analytical opportunities.

2. Theoretical Perspectives

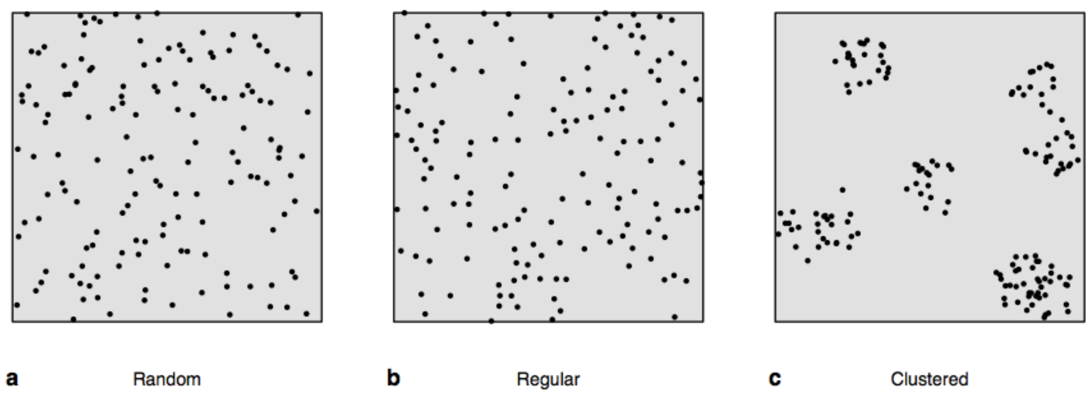
2.1 Point-based Simplifications

A dot on a map is usually a considerable simplification. Whether our concern is about the proper 2- or 3-dimensional representation of a real world entity, its more complicated expression in space-time, or the possible webs of cultural meaning that might envelope it, we certainly risk much by this kind of spatial abstraction. There are also further trade-offs to do with how we record such points, between time and effort on the one hand, and any possible archaeological insights we might derive on the other. Do we require great spatial accuracy (such that measured coordinates are close to the actual absolute values), great spatial precision (where measurements of a given location are highly reproducible, but not necessarily accurate) or some combination of the two? Furthermore, points can also be thought of as highly simplified events in space-time: as such, they often involve only a fuzzy or very loose sense of duration (how long an event lasts for) and equivalence (to what extent an entity in one time step can still be considered the same entity in the next). Indeed, if we observe such events in traditional blocks of archaeological time (e.g. periods or phases), this lumping procedure is a further abstraction with its own additional methodological implications and risks.

An important initial stage of spatial (and spatio-temporal) analysis therefore involves deciding what kinds of simplification and trade-off are acceptable for what applications, as well as how best to make use of the information we already have.

2.1 Spatial Randomness, Regularity and Clustering

Assuming for a moment that point-based abstractions are sometimes justifiable, what do we then want to say about such spatial distributions? The main formal question in the past has been the degree to which a distribution departs or not from what we might expect if we simply scattered points at random across the study area. The latter random, purely 'stochastic', process establishes a theoretical baseline usually referred to as 'complete spatial randomness' or CSR. Spatial statisticians also tend to assume that the underlying process responsible for generating this random point pattern operates in roughly the same way across the whole study area (i.e. it is 'homogeneous' and 'stationary'[1]), and that if we were to consider the number of points falling in each of a series of similar sub-units across this area, we would find that their densities (the more common technical term is point 'intensity', as used hereafter) follow a Poisson distribution and are said to be a realisation of a spatial Poisson process. Figure 1a depicts an example of a random distribution of points in a rectangular area, generated according to a Poisson process.



[1] In certain contexts, there are differences between these two terms, but for economy in the discussion below, they are used interchangeably, as are their opposites (inhomogeneity, heterogeneity and non-stationarity, see below). One related aspect of spatial data that does not receive any attention here, but which implies a limited form of non-stationarity, is anisotropy (i.e. situations in which points are found more frequently in certain prevailing directions; see Markofsky and Bevan 2011 for archaeological discussion).

Figure 1. Hypothetical examples of random, regular and clustered point patterns.

In contrast, Figures 1b-c depict two alternative patterns in which the point distribution is (b) more 'regular' (also often described as 'dispersed') or (c) more 'clustered' (also often described as 'clumped' or 'aggregated'). Sometimes such patterns are intuitively obvious and we could get away without using statistics to consider them, but often our spatial intuition is misleading: for example, some people would suspect slight clustering in the figure 1a (in fact, it is purely random), while others might not suspect regularity in figure 1b (in fact, there is an arbitrarily imposed minimum distance between points). It has therefore long been acknowledged that the role of quantitative spatial analysis is partly: (a) to arbitrate in situations where spatial patterns are uncertain, (b) to characterise such relationships in ways that are useful for explicit comparison, and (c) to offer a formal platform for suggesting possible processes and behaviours behind such spatial patterns. In particular, we often assume that, behind any patterns of regularity or clustering are also some interesting alternative processes, beyond one that is purely stochastic and Poisson. Regular patterns are often thought to be the result of 'inhibition' processes. For example, for human settlements we might think of the way in which the existence of one settlement might inhibit the creation of another one immediately next to it (e.g. because of competition over resources, see below). For artefact distributions, regularity can be generated by various kinds of post-depositional, taphonomic sorting or due to very deliberate human decisions about artefact placement. Clustered patterns, in contrast, are often the result of 'attraction' processes. We might think of the movement of people towards larger settlements because of a variety of the advantages such aggregated locations might offer. For artefacts, we can think of processes of discard and subsequent breakage *in situ* that encourage very clumped scatters of such finds in the archaeological record.

2.3 Spatial Inhomogeneity

So any point pattern documented across a given study area (of whatever archaeological size, from one observed under the microscope, to one found on a house floor, to one seen across a whole landscape) can be thought of as a realisation or one or more underlying processes (see also O'Sullivan and Unwin 2003: 51-75). In the simplest null case, a single random Poisson process is involved. In other, still simple cases, a non-random process is at work, but only one, with effects that are homogeneous across the entire study area (even if the pattern manifests differently at different spatial scales, see below). However, in many real world examples, it is likely that multiple processes are at work and/or that they behave differently in different parts of the study area (i.e. they can be described as 'inhomogeneous', 'heterogeneous' and/or 'non-stationary'; for an archaeological example with aggregated count data rather than point patterns, see Bevan and Conolly 2009). Given the prevalence of inhomogeneous distributions in real life, it is both theoretically and practically useful to distinguish between the 'first-order' and 'second-order' characteristics of a given point pattern (e.g. Bailey and Gatrell 1995: 32-5). First-order characteristics are those that describe the average intensity of points across a given region (if this average intensity varies spatially then the point pattern can be called inhomogeneous), and first-order effects refer to one or more external processes or phenomena that encourage the intensity of points in the study region to vary at different locations. In contrast, the second-order characteristics of a point pattern describe the relative intensity of points as influenced by the spatial configuration of other points in the study area (i.e. the pattern's covariance structure), and reflecting different kinds of internal interaction effects among points, such as propensities for attraction or inhibition. A basic lesson from many practical analyses is that it is difficult, and often entirely misleading, to consider second-order effects before properly accounting for first-order effects.

We return to these issues in the first two case studies below. For now it is simply worth noting that, we can easily build simulations in which to observe what kind of spatial pattern is produced by any single realisation of a particular point process of know design. In real world contexts however, the process that generated the points is typically unknown, and the challenge becomes the degree to which we can learn about what the first- and second-order effects might be solely via analysis of the resulting pattern. Ironically, while it is fair to say that many archaeologists would be loosely and informally aware of such complex spatial considerations when it comes to their interpretation of the archaeological record, the formal quantitative tools they have so far used have been stuck in some rather idealised and methodologically-quarantined boxes. For example, at the scale of landscapes and archaeological sites, 'predictive modelling' (e.g. Mehrer and Wescott 2006; Verhagen and Whitley 2011) [2] has been a commonplace way of assessing first-order properties, demonstrating, for example, correlations between the probability of discovering sites in a particular study area and the distribution of one or more environmental variables (e.g. soils, slope steepness, access to water, etc.). Conversely, nearest neighbour tests and quadrat counts have typically been used to assess site spacings, with the implicit assumption that second-order interaction effects are often at work. However, rarely if ever, are these two methods brought together to treat the issue as an analytically-related whole.

One final, complicating factor for archaeologists is the fact that archaeological observations are very partial, imperfect records of past activity. Much of the variability in our observed spatial patterns in archaeology is due to patchy levels of archaeological preservation and investigation. For example, most site distribution maps are the result of historically-complex sites and monuments records or unsystematic surveys – many of the perceived clusters of observations are to do with where people have recently looked, where modern development has recently exposed new archaeology, etc. These issues can also be conceived of as kinds of first-order variation in intensity, but ideally we would want to distinguish them from taphonomic and human behavioural effects in the past, ultimately so that we can offer some useful archaeological interpretation.

3. Multi-scalar and Monte Carlo Approaches

[2] In archaeology, this term has developed an unnecessarily narrow meaning, related to cultural resource management and models of site location probabilities. However, beyond archaeology, it is just a general term for any kind of model that leads to explicit predictions of one kind or another.

This section moves from general theoretical considerations to explore the relevance of a variety of recent methods for characterising point patterns and processes. It begins with a hypothetical example to fix some ideas before considering a real, intra-site case study.

3.1. A Hypothetical Example

It may sometimes be very difficult to wholly separate complex first- and second-order effects (some operating in the past, some in the present) in many archaeological datasets, but there remain many advantages to conceptualising point distributions this way. Alongside the theoretical issues raised above, a whole host of more advanced analytical methods have been discussed in the spatial statistical literature over the last 20-30 years (for a recent overview: Gelfand et al. 2010: 263-423). These have been used in certain applied fields such as astronomy or ecology for a long time, but have been slow to percolate into other disciplines such as archaeology.

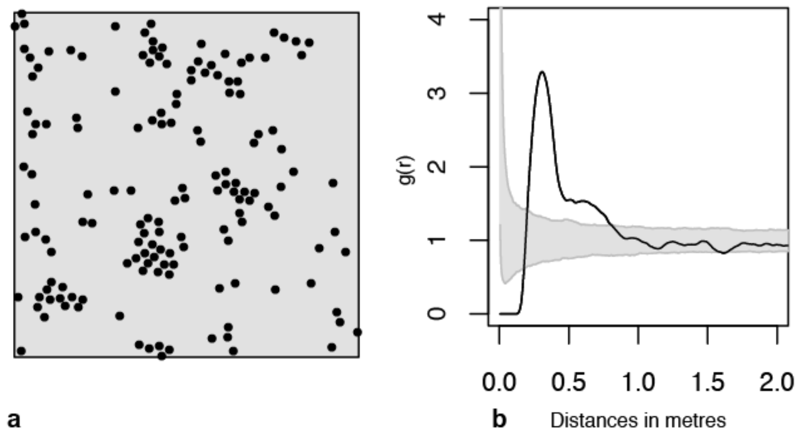


Figure 2. Multi-scalar Monte Carlo methods: (a) a point pattern, in a in a notional 10x10m study area, that is both regular and clustered at different distances; (b) a pair correlation function identifying a switch from significant regularity to significant clustering at c.0.25m (with the latter slowly tailing off thereafter). The grey are in b) is a 95% envelope based on 999 random simulations.

Two key methodological advances over the last 35 years (outside of archaeology) have been methods that: (a) deliberately seek to address point pattern and process at several different spatial scales, and (b) employ a family of randomisation tests know as Monte Carlo simulation (Robert and Casella 2004), which leverage the speed of modern computational platforms to provide a powerful and flexible way of testing spatial patterns, particularly in cases where the study area is irregular or the underlying effects are complicated (for other archaeological applications of Monte Carlo simulation, see Fisher et al. 1997; Drennan and Peterson 2004; Crema et al. 2010). Figure 2a, for example, presents a toy example of a point pattern produced by a known process (see Lennard-Jones 1924) in a notional 10x10m study area which leads to (a) a strong tendency for regular spacing over very short distances (up to 0.25m in this case), but thereafter also (b) a further tendency for clustering at medium distances which gradually tails off to a random pattern at larger ones. In this case, no first-order effects are present and the process operates in a uniform way across the whole hypothetical study area. To what extent however, can we find methods that correctly identify these different scales of second-order effect based solely on analysis of the resulting site distribution? If, for example, we calculate a traditional nearest neighbour index (Clark and Evans 1954; Hodder and Orton 1976: 38-51) that has been, for better or worse, the bread-and-butter of archaeological point pattern analysis for many years, it misleadingly suggests that the pattern is random or only very slightly clustered (r=0.91).

There are however more recent spatial statistical methods that consider multiple scales of second-order patterning and explore how likely or unlikely they are to have occurred by chance (for a technical overview, see Gelfand et al. 2010: 263-423). Perhaps the most common of these is the K function and its more readable, slightly transformed version, the L function (originally Ripley 1977; and for some exploratory archaeological uses, Orton 2004; Bevan and Conolly 2006; Vanzetti et al. 2010). Here we emphasise another related method, the pair correlation function (PCF), which is less well known, but arguably more useful in many circumstances (several similar functions go by other: see Ilian et al. 2010: 218-23; Wiegand and Maloney 2004, Perry et al. 2006). A PCF measures the intensity of points in donut-shaped rings (annuli) around each point and, as such, is not a cumulative statistic in the same way as a K or L function (the latter two effectively measure the intensity of points in ever expanding circles that include all previous, smaller ones).

Figure 2b shows PCF results for the simulated point pattern. The x-axis measures the separation distance between points and the observed results are presented as a black line. This observed result begins well below the theoretically random threshold of y=1, indicating the possible regularity of this pattern at short distances, then climbs well above this threshold, indicating medium distance clustering, before dropping slowly back down towards y=1. For a variety of reasons however, this theoretical y=1 threshold is often an unreliable baseline, and it is more useful to use Monte Carlo methods that offer an 'envelope' of possible values that we might expect under a null model in which the point process generating this pattern is assumed to be wholly random. This is done by repeatedly generating sets of an equivalent number of random points, and then plotting maximal and minimal PCF values at each distance range. In the case of figure 2b, the grey shaded area marks out, not the full range of random PCF values, but an envelope enclosing the middle 95% of PCF values from 999 simulation runs.[3] Where the real, observed values are larger than this envelope, the observed pattern can be considered clustered at that distance, whereas where they fell below the envelope, they are more likely to be regularly spaced. In this example, the PCF successfully and accurately documents the shift from

[3] For simplicity and consistency in each of the analyses developed in this paper, we have run 999 Monte Carlo simulations, have combined these with the observed values and have then taken the 25th and 975th ranked values to define the borders of the envelope depicted in each plot. At first glance, it might seem as if these envelopes could be treated as also defining a 0.05 significance level, but in fact this is potentially misleading for tests that consider multiple critical values simultaneously. Alternative envelope calculations that do produce exact significance envelopes are feasible (e.g. see ?envelope in the R spatstat package), but are more complicated to implement consistently across the different methods used here, so they have not been included.

significantly regular to clustered effects (the critical feature in this kind of plots is usually the point of inflection at about y=1, e.g. here at ca. 0.25m, rather than the top or bottom of observed humps in the PCF), and thereafter the slow tailing off of this clustering until the pattern becomes wholly random. In fact, the non-cumulative nature of the PCF offers certain advantages over K or L functions for analysing patterns with these kinds of multiple scales and different kinds of interaction, although, in general, such methods offer complementary perspectives.

3.2. Crossbow Triggers and Qin Terracotta Warriors

Of course, real archaeological distributions rarely, if ever, manifest themselves as such completely recovered, simply bounded datasets. The first of our three archaeological case studies therefore explores some of these analytical issues as they arise at the intra-site scale. The tomb complex of the first Chinese emperor, Qin Shihuang (259-210 BC), is famous, amongst other things, for its pits of life-sized, terracotta warriors, buried in battle formation, with full military equipment. As an example, we can consider the distribution of bronze crossbow triggers (this being the only part of the crossbow that survives archaeologically) that were found alongside the warriors in the easternmost part of pit 1. A plot of these artefacts against the warriors (Figure 3) makes it clear that the overall crossbow trigger distribution itself is strongly clustered in space due both to the shape of the corridors and the nature of terracotta army's battle formation (with crossbowmen only in certain parts of the army, particularly along the flanks). This, in and of itself, is probably not something we need to assess via more complicated statistical treatment, but in passing, it is worth noting that the corridors represent a very irregular study area that raises some complicated issues to do with 'edge effects' (e.g. how we handle the fact that the annuli must be truncated to reflect the fact that points cannot fall beyond the corridor area, as well as the possible inaccuracies that arise from our lack of knowledge about areas immediately west of the excavated portion of pit 1).

For our purposes here, however, a key interest is not how to characterise the overall pattern of triggers, but how we might modify the methods introduced above to consider patterning amongst different sub-groups of crossbow trigger. More precisely, metrical, typological and materials analysis of the triggers has been able to distinguish subtle but undeniably different trigger sub-groups that suggest the existence of different weapon-casting moulds, different metallurgical workshops and/or different organisational practices. The difficult question therefore becomes: how do we assess the spatial distribution of the trigger sub-groups while controlling for the overriding spatial structure of trigger distribution in general?

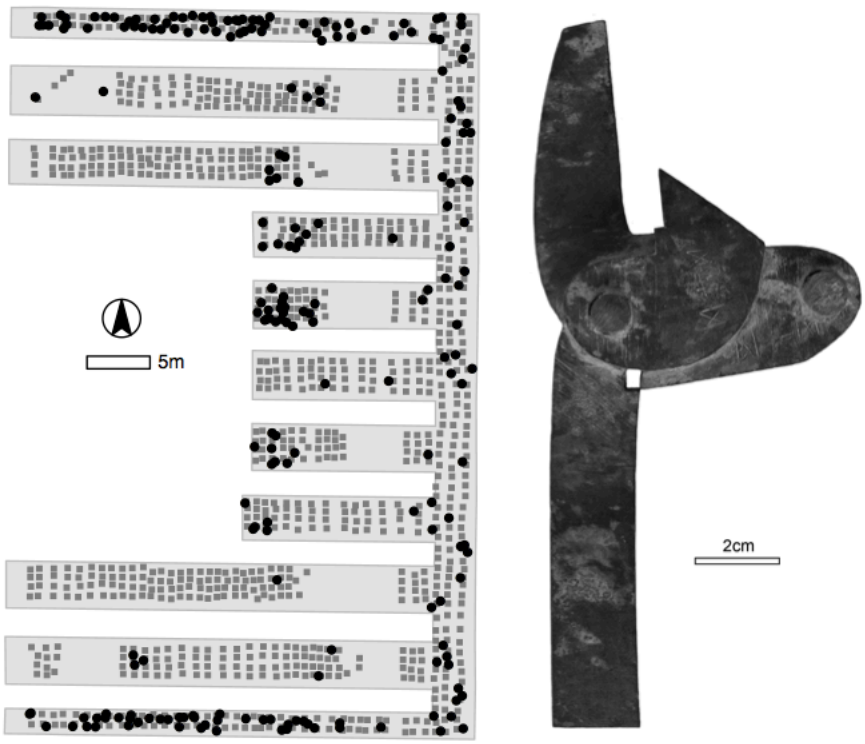


Figure 3. Intra-site spatial distributions: (a) Qin terracotta warriors (grey squares) and bronze crossbow triggers (black circles) in the easternmost parts of pit 1, (b) a photograph of a bronze cross-bow trigger.

A good example is the trigger sub-group shown in figure 4a. This is a group that, when studied in detailed, exhibits small but distinct morphological and typological differences from other triggers. To recap, when we focus on the possible spatial patterning of this sub-group, we clearly want to control for the spatial structure of the triggers as a whole (and by extension the formations of crossbowmen). To do so, we run a Monte Carlo simulation in which the triggers attributed to this particular sub-group are assigned at random amongst the overall trigger assemblage. In fact, the group 2 triggers in the pit are, themselves, visibly clustered, beyond the pattern imposed by the battle formation (Figure 4a) and, again, there may not be a need for a formal method to recognise it in this case. However, it is useful to consider this particularly clear-cut example as a proof of concept, and in the knowledge that such standardised evaluation will be far more important in other less obvious cases. Figure 4b shows a pair correlation function in which this clustering is very evident in the observed result substantial deviation above the 95% envelope. More precisely, the plot indicates particularly strong clustering of this sub-group up to distances of perhaps 3-4m radius and then up to 7-8m, with further possible clustering at much larger distances. There are some interesting processes that are likely to be behind such clustered patterns of trigger sub-groups in the pit. For example, they may reflect different workshops producing marginally different crossbow triggers and procedures for the storage and placement of the crossbows in the pit in batches (e.g. zones of the pit that were equipped with crossbows in one go). Applied more broadly to other trigger sub-groups, other weapon types and other artefacts in the pit, such analyses can begin to map out coherent activity spaces and explore how consistent they were in size, arrangement etc. (see especially, Li 2012; Martinón-Torres et al. in press).

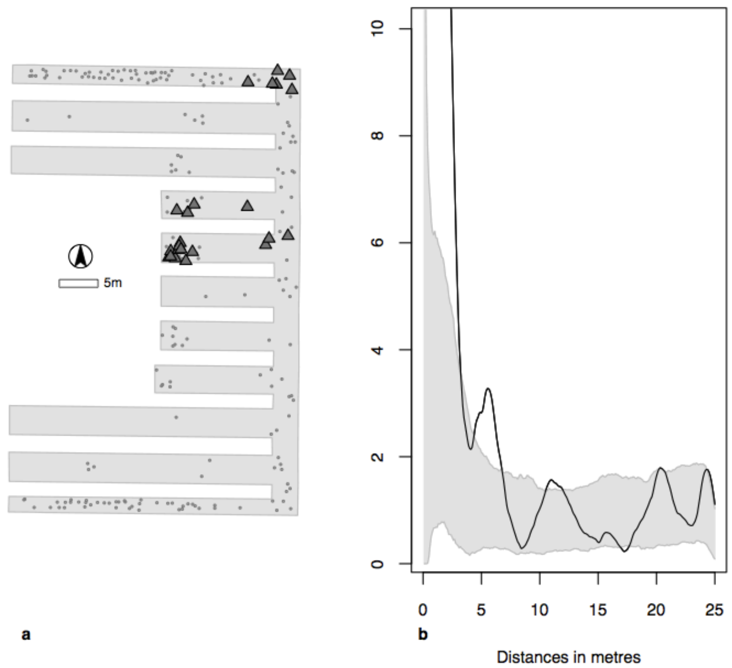


Figure 4. Spatial analysis of trigger groups: (a) group 2 triggers shown as triangles and the others as grey dots; (b) a pair correlation function (observed values in black and 95% critical envelope in grey)

4. Inhomogeneous Point Process Models

The above case study demonstrates, via a deliberately straightforward example, that such methods can formalise our assessment of spatial patterns at multiple scales, even in the presence of other confounding spatial factors (such as the shape of the corridors and the clustering of crossbow triggers as a whole). In their original form, methods such as K, L or pair correlation functions were not easily applied to these kinds of inhomogeneous and edge-affected cases, but such problems are now becoming increasingly tractable. The second case study considered here, explores the potential of such inhomogeneous approaches for assessments of site location at the landscape scale. It considers some Iron Age I (ca. 12th-11th centuries BC) settlements documented by fairly systematic surface survey in the central part of the West Bank (modern-day Israel and Palestinian Territories, for the survey, see Finkelstein and Magen 1993; Finkelstein and Lederman 1997). In particular, we focus on an area of hilly dolomite upland of some 766 sq.km, across which 99 sites of this period have been documented (Figure 5a). This choice of area is deliberate: it reduces the range of complicating factors that need to be considered below, both because it was investigated in a fairly even way by a single archaeological project, and because it covers an area of generally consistent underlying geology (with some knock-on implications for soils, topography, etc).

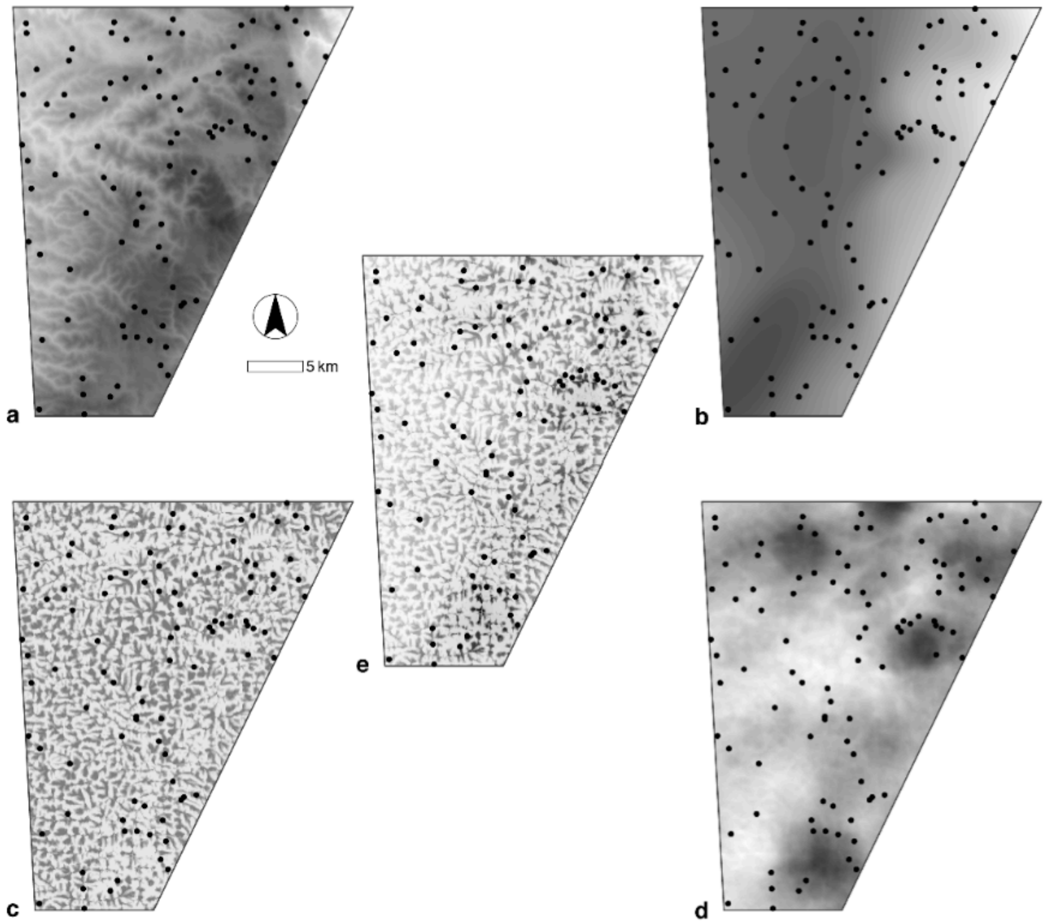


Figure 5. Iron Age I sites in the central West Bank and four possible first-order covariates: (a) elevation (light to dark ranges from 135-1010m ASL), (b) average annual rainfall (dark to light ranges from c.335-720mm), (c) ridge landforms (darker is more likely to be geomphorphometrically classified as a ridge), (d) topographic wetness index summed over a local neighbourhood (darker is wetter), and (e) a prediction surface based on the three significant covariates (darker is higher point intensity).

A quick visual inspection of figure 5a suggests informally both that there might be a first-order trend towards slightly greater densities of settlement at higher elevations, and also that there might conceivably some regular spacing to some of the settlements. We can therefore build some formal point process models to consider whether environmental affordances such as elevation are indeed significant, and above and beyond this, whether there is yet a further second-order propensity for the location of one settlement to inhibit the location of another nearby. We begin by considering, as examples, four related environmental affordances – elevation, average annual rainfall, ridge-top landforms, and topographic wetness in a local catchment (Figures 5a-d).[4] This

[4] The digital elevation model (DEM) used here is NASA's 90m SRTM dataset (Jarvis et al. 2008). The rainfall data has interpolated from 50mm contours of average annual precipitation (the original contours are courtesy of the GIS Center, Hebrew University of Jerusalem). Ridge-like landforms were defined from the DEM via a fuzzy feature classification across focal filter scales from 3x3 to 11x11 cells (Fisher et al. 2004). Catchment-based topographic wetness was calculated via focal filtering of a standard topographic wetness index surface (itself derived from the DEM) in a way that summed all values within a circular neighbourhood of 2.5km radius (about half an hour's walk and a common threshold for daily travel budgets).

selection is prompted in part by many commentators' informal impressions that rugged topography and hydrology were important factors behind settlement locations in this region and period, for a variety of practical reasons (e.g. Zertal 1988; Gibson 2001; in fact many other possible covariates have been explored but are not considered here). Univariate regression of binned versions of each of these variables against site intensity (figures 6a-d) suggests that rainfall is not a particularly good predictor of the intensity of sites across the landscape, but that the other three variables have significant positive correlations ($p<0.05$ or better). In other words, sites are more common a) at higher elevations, b) where the landforms are more ridge-like in shape, and c) where patches of ground offer better access to surface water or soil moisture. Informally, it appears that ridge-like locations might be the most influential of these.

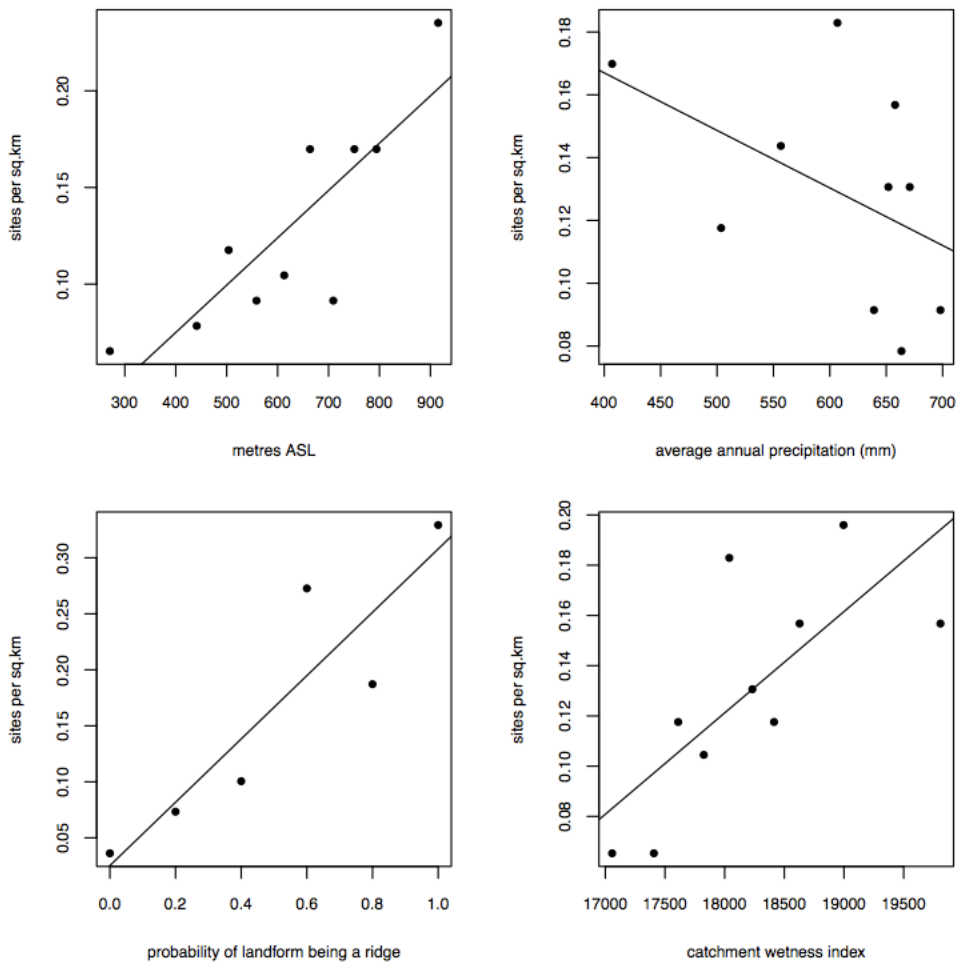


Figure 6. Univariate correlations between site intensity and a) elevation above sea-level ($r^2=0.72$), (b) average annual rainfall ($r^2=0.23$), (c) ridge landforms ($r^2=0.82$), and (d) topographic wetness index summed over a local neighbourhood ($r^2=0.54$). The intensities for a, b and d have been summarised in decile bins of the covariate, while for c, the x-axis probabilities are discontinuous due to the nature of the fuzzy geomorphometric classification used.

These univariate regressions provide a guide to likely relationships between site intensity and various first order effects (note, as above, that we have already removed other possible first order effects by choosing a study area that is broadly one type of geology, and that has been surveyed by one field project with fairly consistent methods). If we now run a multivariate regression and select the best combination of these four variables via stepwise comparison (minimising an Akaike Information Criterion), we find that rainfall is excluded as we might expect, that the other three variables are all significant ($p<0.05$ or better), and that this new model with a first-order trend is substantially more effective that a null, random hypothesis. We can then create a predicted first-order intensity surface (figure 5e) that can be used to return to the question of second-order interactions in a more complete way.

First, as a point of contrast, it is worth considering a pair correlation function of the settlements sites, along with an envelope of wholly random Monte Carlo simulations (figure 7a). The observed PCF shows something very close to regularity at shorter distances of up to ca.1km, although the Monte Carlo envelope suggests that this might be of marginal significance. If we then look at a simple histogram of nearest neighbour distances (figure 7b), we can see a spike at just over 1,000-1,250m and a Monte Carlo 95% envelope suggests that this pattern is unlikely to occur by chance (for this method with histograms of nearest neighbour distances, see also Wilson and Melnick 1990). In other words, the nearest neighbour histogram provides a slightly more discerning picture of very short distance patterning that confirms the evidence for regular spacing that was initially visible in the PCF (the same observation is valid for other multi-scalar functions such as K and L as well). At this stage however, we cannot be sure whether such regular spacing has been induced by the spatial structure of some important external influence on site location (e.g. evenly spaced ridge-tops) or is due to internal processes that inhibited settlements being located close to one another (e.g. competition over resources).

Figure 7c seeks to tease apart the relative contribution of first- and second-order effects by showing the same nearest neighbour histogram for the observed values, but this time with a simulated envelope conditioned on the spatially inhomogeneous intensities predicted by our first-order covariates. Put plainly, the Monte Carlo sets of points are now forced to respect the spatial inhomogeneity modelled by the predicted intensity surface (figure 5e). As mentioned above, we might conceivably anticipate that one or more of the first order variables, such as evenly-spaced ridgelines, might have accounted for some of the short distance regularity in settlement, but this is not the case. Of course, it is possible that this continuing regularity simply means that some important environmental covariate has not been considered, but nevertheless, we are now at least moving closer towards ruling this possibility out.

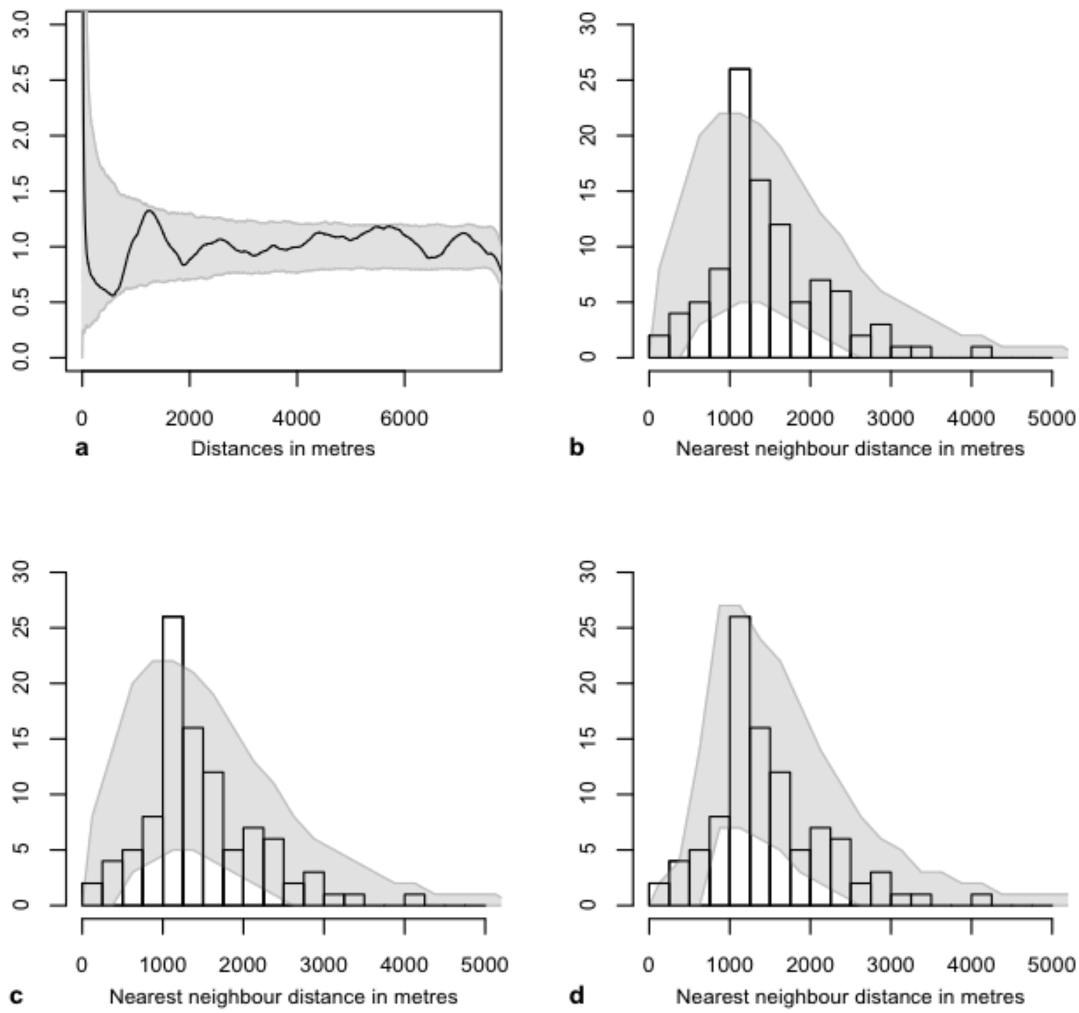


Figure 7. Point process models and goodness-of-fit: (a) a pair correlation function of the observed sites with a 95% envelope from wholly random Poisson process, (b) a histogram of nearest neighbour distances with a 95% envelope from wholly random Poisson process, (c) a histogram of nearest neighbour distances with a 95% envelope also conditioned on the first-order covariates model, (d) a histogram of nearest neighbour distances with a 95% envelope also conditioned on both the first-order covariates and a second-order, area-interaction model.

Finally therefore, we can explore the goodness-of-fit (via informal visual comparison here, though more formal statistical treatment is also possible) of an explicit model of what might be causing these second-order point interactions. Perhaps the most relevant one is Baddeley and van Lieshout's "area-interaction model" (1995) that generates patterns of inhibition and clustering with reference to a defined circular neighbourhood around each point. The implicit idea of this model – that points have notional territories of influence around them – is obviously attractive given our understanding of how many human settlements work. We can draw upon our knowledge of the observed spacing between settlements and set the parameters to suggest a radius for the interaction neighbourhood of just over 655m radius (half the median nearest neighbour distance), and inhibitive effects that are very strong but not absolute within this zone. These parameters lead to sites spacings that are often twice the neighbourhood radius and which often suggest formal or informal village catchments of less than 135 ha, with such a scale being not unreasonable given evidence for fairly small Iron I community sizes of dozens to no more that a couple of hundred people in this area. Figures 7d demonstrates that the histogram of observed nearest neighbour distances now falls within the Monte Carlo envelope. As we discuss below, other explanatory models might conceivably offer better or equivalent fits, but by narrowing down the range of possibilities in this formal manner, we clarify our thinking about what might be plausible kinds of causal phenomena in a very useful way.

5. Models with Temporal Uncertainty

Our final case study is again one focused on settlements and landscapes, but with a greater emphasis on diachronic comparison in the presence of uncertain dating. Temporal uncertainty is an elephant in the room of much archaeological interpretation. It is a near ubiquitous feature of archaeological datasets, whether these are radiocarbon dates, geoarchaeological deposits or individual artefacts. There is insufficient room to discuss this topic at length, but for our purposes here, one primary risk in the spatial analysis of point distributions is that they might reflect a chronological palimpsest that thwarts our ability to unpick single-period, contemporary point patterns. This is also a topic to which many archaeologists have discretely turned a blind eye: for example, it is no more than a convenient analytical assumption that all 99 of the Iron I settlements in the central West Bank study area discussed above were inhabited at exactly the same time during that phase which spans a couple of centuries (indeed a few of the unusually clumped sites in the eastern part of the study area might be seasonally occupied). In other cases, the chronological range of the sites under investigation is even broader and the risk of drawing misleading conclusions is correspondingly greater. This is especially true with regard to the assessment of second order effects and the processes that lead to them. Michael Barton (this volume) nicely outlines an example of a regularly spaced pattern of small sites in north-central Arizona that might be due to patterns of shifting clearance, short-term cultivation and abandonment, in which many of the sites involved belonged to the same broad period, but might not be strictly contemporary. The implied processes of interaction are, in this case, quite different and the discussion at the end returns to the well-known problems of equi-final models that this raises.

5.1 *Aoristic Methods*

One way to engage more effectively with temporal uncertainty is for us to make the best of all our available temporal information, however fuzzy. Occasionally, we can define an explicit probability distribution that suggests how likely it is for an event is to have occurred at a certain stage in time based on a range of sources of knowledge (e.g. diagnostic artifacts, clear stratigraphic relationships, absolute radiocarbon dates, etc.). Even so, such results rarely produce a normal distribution that can be conveniently summarised by a single summary value and confidence interval. Instead, we are more likely to have the kinds of irregular probability distribution often produced, for example, via Bayesian modeling of calibrated radiocarbon dates and associated soft information. More importantly, such information-rich cases are rare: in most instances, we can only suggest a very approximately bounded ‘time-span’ within which the event is likely to have occurred and, within this, assume a simple, uniform probability distribution (i.e. implying that an event has a similar chance of having occurred at any stage within the time-span). 'Aoristic' analysis is an approach that provides a way of quantifying these temporal uncertainties and incorporating them into subsequent analysis (particularly in the case of simple timespans where we assume uniform probablities). It was initially developed in criminology (Ratcliffe 2000), and has subsequently also been adopted by some archaeologists for looking at both individual artefacts and larger archaeological sites (Johnson 2004, Crema et al. 2010; Pentedeka et al. 2010).

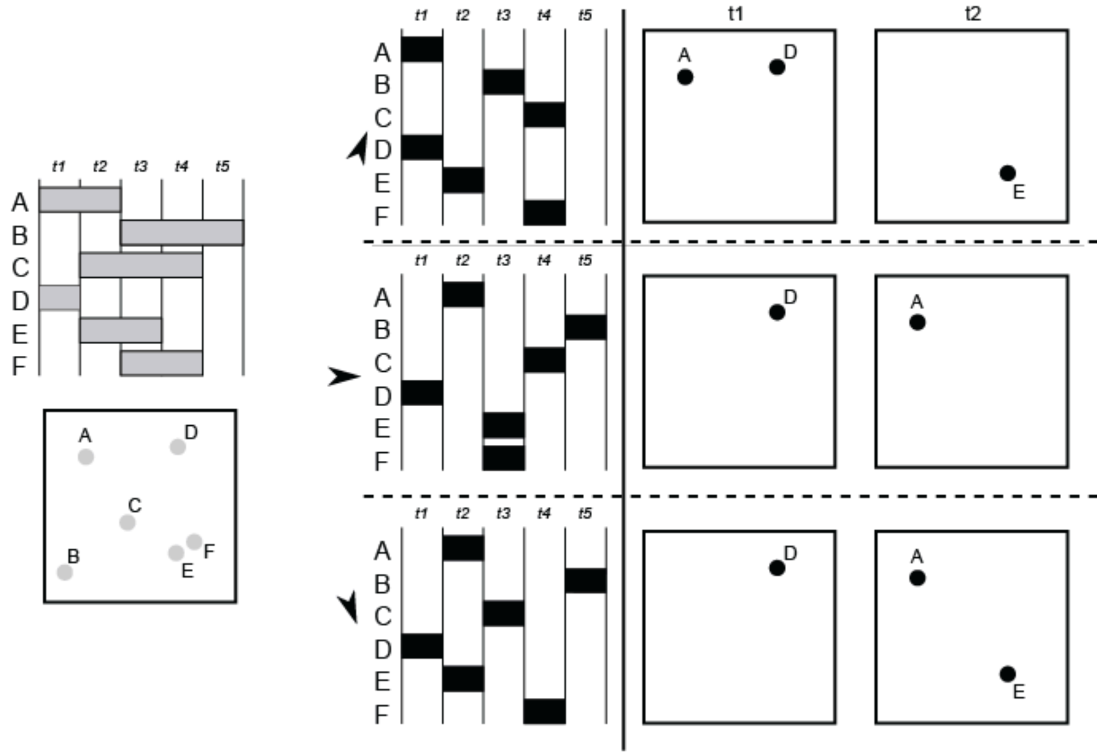


Figure 8: Temporal uncertainty in point patterns. The left panel depicts a simple hypothetical point pattern and (above this) our degree of temporal knowledge about each of the six point events (A-F) over five time-blocks (t1-t5). The grey horizontal bars represent the time-span of each event, showing that point event D has low uncertainty (the point event belongs exclusively to time-block t1) and event B has higher uncertainty (the point event belongs somewhere between time-blocks t3 and t5). The three panels in the the middle show three possible realisations of the actual temporal (middle panel) and spatio-temporal patterns (right panel, for time-blocks t1 and t2).

More precisely, given a specific set of points with their temporal probability distributions, there will be a limited number of *possible* spatio-temporal patterns that might actually have arisen. Instead of ignoring this uncertainty and producing a single, but misleading, spatial analysis, we can generate different possible spatial patterns based on these temporal probabilities and then obtain a distribution of the more and less likely results. Figure 8 is a schematic representation of both the problem and the possible solution: take, in this case, six point events (A-F) that each occurred in one of five time-steps (t1-t5), but can often only be ascribed archaeologically to wider time-spans (i.e. figure 8, left). Each of the actual scenarios in the middle and right hand-panels of this figure are possible realizations of the pattern, amongst many others, given the state of our temporal knowledge. The only way to explore what possible spatial patterns might really have been present is therefore to analyse a whole host of possible realisations and explore if there are any first- or second-order spatial properties that persistently crop up. If for instance, 90% of all the possible point patterns are spatially clustered, we will have a relatively high confidence that the observed pattern was indeed clustered.

However, while from a theoretical standpoint, it might be tempting to consider each and every possible spatio-temporal configuration, in practical terms this is computationally prohibitive as the number of possible scenarios is often intractable. The alternative however is simply to sample a finite number of possible realisations via Monte Carlo simulation and calculate the frequency of certain spatial patterns in the results (see Crema et al. 2010 for further details; and Izquierdo et al. 2009 for a similar perspective).

5.2 Middle Jomon Settlement

Our final case study combines aoristic and Monte Carlo methods to consider settlement patterns amongst the Jomon hunter-gatherers of central Japan. In a sense, it offers an ideal case for tackling the issue of temporal uncertainty because, while careful pottery study and an amazingly dense number of emergency excavations (Habu 2004; Kobayashi 2008) provides one some of the most detailed spatial distributions and relative chronologies known for any prehistoric complex hunter-gather groups worldwide, it remains true that some Jomon pithouses and broader settlements can be ascribed to only fairly broad chronological rnages whilst others can be dated far more accurately.

Moreover, Jomon settlement patterns exhibit some interesting possible patterns that may relate to changing demography, social practices and subsistence strategies. For example, several studies (e.g. Imamura 1996) have indicated a sudden rise in the overall number of Jomon residential units during the first part of the Middle Jomon period (ca.3530-2470 cal. BC), followed by a rapid collapse after few centuries. Some authors explain such dynamics as due to increasingly intensive use of certain plant resources during the early Middle Jomon that made it possible to maintain higher population densities, but which became more problematic during a subsequent climatic cooling phase in late Middle Jomon which may have led to a reduction in the overall availability of these resources (Imamura 2002, Habu 2008). In terms of spatial patterning, most commentators agree that there were larger, more nucleated settlements prior to the proposed population collapse and more dispersed, smaller settlements after it. While such a broad dichotomy seems plausible for the Middle Jomon, it remains difficult to consider tempos of change over smaller timescales or to compare these processes with those in earlier or later periods of Japanese prehistory. Aoristic analysis and Monte Carlo simulation can however provide a good analytical framework for such research.

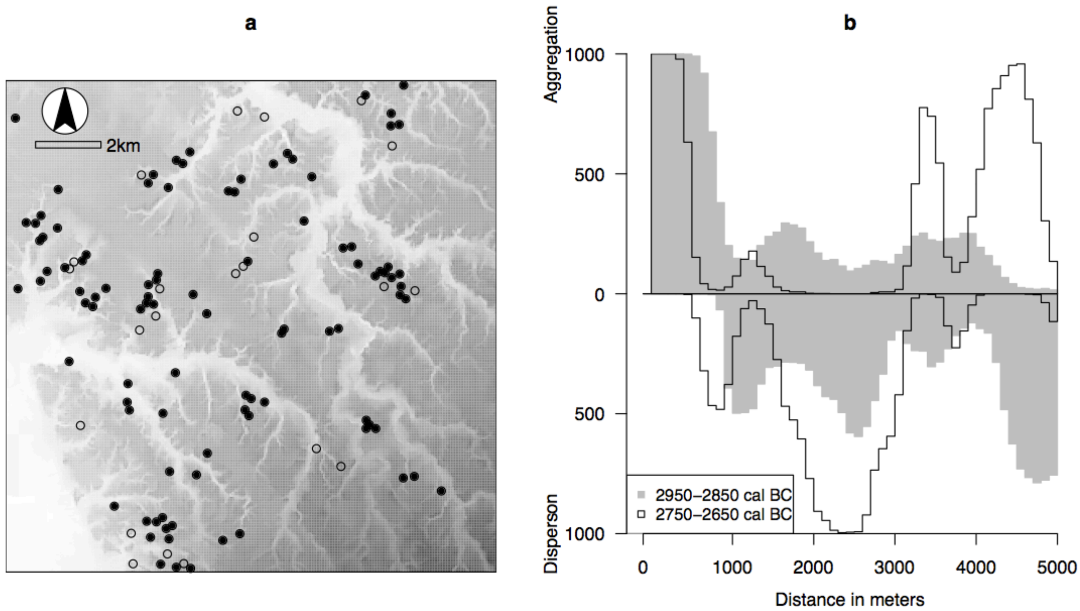


Figure 9: a) Distribution of excavation units containing at least one pithouse attributed to Early to Late Jomon (filled and hollow circles) and locations where at least one pithouse with a probability of existence higher zero at one of the two temporal blocks examined is present (filled circles). b) Number of PCFs with significant clustering (above the horizontal line) and significant dispersion (below the horizontal line) for time blocks 2950-2850 cal. BC (grey shaded bars) and 2750-2650 cal. BC (hollow bars) with 1000 simulated spatio-temporal patterns.

The case study area for the analysis that follows is located on the western side of Tokyo Bay, near the modern city of Chiba. We have chosen an arbitrary square-shaped area of 15x15 km within which 120 separate open-area excavations have documented some 1418 Jomon pithouses that can each be attributed to somwhere between the Early and Late Jomon period (ca. 5050-1270 cal. BC; figure 8a). For each of these pithouses, we can define a more precise time-span of existence from the description of pottery and associated artefacts available from excavation reports. We can then spit the whole 5050 to 1250 cal. BC timeframe into arbitrary chronological blocks of 100 years each and calculate the probability that a given pithouse actually exists during that block. For our purposes here, we compare two distinct time blocks – 2950-2850 cal. BC during a phase of population increase, and 2750-2650 cal. BC at the observed peak in population. For each of these two, we then generate 1000 Monte-Carlo simulations of possible spatio-temporal patterns. In these simulations, and again for the purposes of this example, we designate a new settlement to exist at a certain location, if an excavated area at that spot is allocated at least one pithouse for that chronological block.

The resulting simulations provide a series of realized settlement patterns: in order to explore differences in the respective spatial distributions for each period, we computed a series of PCFs. Just as in our earlier case studies, each observed PCF can be compared to an envelope generated from 999 Monte Carlo sets, each with an identical number of points and each only allowed to exist within the 120 parent locations where Early to Late Jomon pithouses has been actually been recovered. This method of constrained randomisation is similar to the one adopted for the case study on bronze crossbow triggers, and allows us to account both for the patchy nature of modern excavation and for some general first-order, locational choices the Jomon may have had over the long term (the latter being of great general interest of course, but not for our analysis here).

Figure 9b shows the frequency of instances among the 1000 simulated spatio-temporal patterns where the observed results depart from the 95% envelope of random values. During the earlier phase of population increase (2950-2850 cal BC) settlements appear to be aggregated to varying degrees over distances up to 600 or even 1,000m radius (and indeed, whether we see this as an observation about inter-settlement clustering or simply large extensive settlement areas is partly just a question of semantics). At much larger spatial scales, there are a few instances of dispersed patterns, but these are fairly rare, with an exception at ca. 4,700-4,900m, where about 70% of the simulated patterns showed dispersion. Also notice how, at some spatial scales, (e.g. at ca.1,500m), the number of instances of clustering and dispersion are roughly equal, implying that our levels of available information are insufficient to draw any robust conclusions. A few hundred years later, when the population size reached its peak (2750-2650 cal BC), the settlement pattern is notably different. Firstly, the short-distance clustering is still present, but now only occurs of a much smaller range (<300m radius). This indicates that clusters of nucleated settlements may have started to decrease in their sizes, an idea that is supported also by a small peak of dispersed patterns at ca. 1,000m. Secondly there is strong evidence for patterns of dispersion (i.e. regular-spacing) at separation distances of c.2,500 meters. Thirdly at ca.4,900 there is a relatively high number of simulations (>95%) that lead to aggregated patterns, the opposite result to the one seen in the previous time-block. Clearly, the patterns for the two different time periods are not the same, with the former being characterized by a greater nucleation and the latter possibly by greater dispersion at medium distances and aggregation at higher distances (i.e. broad clumps of settlement activity with intervening spaces of some 5km between these). While we still need to treat such results cautiously, the analysis suggest that the beginnings of the dispersed patterning know to be present in anger by the mid 3rd millenium BC is already visible during the period of peak settlement, reflecting possible early instances of group fission that have been plasuible argued as being driven by the diminishing availability of local food resources.

6. Discussion

The theoretical discussion and three case studies above should convey the degree to which it is worth reinvesting in the formal modelling of point patterns and processes in archaeology. In any case, from this discussion we can draw out several practical conclusions:

a) Careful definition of a study area is important.
b) Histograms of the distribution of nearest neighbour distances offer a fairly robust way of exploring short-range regularities in point spacing, and can be made more robust as a confirmatory method via Monte Carlo simulation. Traditional Clark and Evans tests are far less discerning.

c) Multi-scalar methods such as K, L and pair correlation functions are potentially useful for understanding second-order interactions, but are inappropriate on their own if there are grounds for thinking the patterns exhibit spatial inhomogeneity.
d) We can use multivariate regression models (in a similar manner to established practices in archaeological predictive modelling) to provide first order measures of the varying intensity of points across a study area as influenced by a range of external variables. These can then offer a platform from which to consider second-order interactions via the above multi-scalar methods even where spatial inhomogeneity in present.
e) The kinds of temporal uncertainty present in most archaeological datasets can be successfully addressed in spatial analysis by adopting a probabilistic and Monte Carlo framework.

It is worth ending this chapter by revisiting one well-known criticism of formal approaches to spatial analysis and modeling in archaeology. A common suggestion is that such efforts are, at best, frustrating and, at worst, have little interpretative value (e.g. Hodder 1977), because: a) several different modelled processes can sometimes be shown to produce the same or similar outcomes (i.e. they are equi-final or convergent) and b) the same model can sometimes be shown to lead to quite different outcomes depending on its exact starting conditions or the role played by random chance (i.e. it is multi-final or divergent). However, while these challenges should encourage us to avoid statements that imply a cast-iron certainty about causal relationships, they should not dissuade us from trying to model them at all: a smaller set of equally plausible models is still better than a situation in which anything goes (see also Premo 2010). In this sense, the above discussion has led us from a traditional quantitative emphasis on simply 'rejecting null hypotheses' towards one in which greater emphasis is placed on comparing the fit of a series of potentially plausible explanations (i.e. in general sympathy with a maximum likelihood approach). In other words, Monte-Carlo methods allow us to embrace equifinality by weighting alternative hypothesis in probabilistic terms. This approach should be welcomed as more in tune with interpretative archaeologies of landscape that often look positively on the co-existence (temporary or otherwise) of rival explanations: there is no reason that a similarly healthy lack of certainty should not be welcomed for quantitative approaches as well.

Acknowledgements
The crossbow trigger case study derives from a much wider collaboration between UCL Institute of Archaeology and the Museum of the Terracotta Army, and our particular thanks to Marcos Martinón-Torres and Thilo Rehren and the Terracotta Army Museum staff. The Iron I settlement survey dataset is one provided by the West Bank and East Jerusalem Archaeological Database (digitallibrary.usc.edu/wbarc/), with slight modifications by Palmisano. We would like to thank Adi Keinan (UCL Institute of Archaeology) for her help with this and related West Bank data, as well as Adi Ben-Nun (GIS Center, Hebrew University of Jerusalem) for allowing us to make use of the rainfall contour dataset. The dataset for the Jomon case study at Chiba has been obtained from the excavation reports kindly made available by the Cultural Properties Centre of Chiba Prefecture Education Foundation. The analysis discussed here was conducted primarily in R (R Core Development Team 2011), with some pre-processing in GRASS GIS (GRASS Development Team 2008) and Landserf (Wood 2009). Our thanks to Adrian Baddeley for advice on point process models and the use of the spatstat package (Baddeley and Turner 2005), as well as to Mark Lake and two anonymous reviewers for comments on a chapter draft.

References

Baddeley, A. J. and van Lieshout, M. N. M. 1995. Area-Interaction Point Processes, *Annals of the Institute of Statistical Mathematics* 47.4: 601-619.

Baddeley, A. J. and Turner, R. 2005. spatstat: An R Package for Analyzing Spatial Point Patterns, Journal of Statistical Software 12.6: 1-41.

Bailey, T. and Gatrell, T., 1995. *Interactive Spatial Data Analysis.* Harlow: Longman.

Bevan, A. n.d. Archaeological Sites and a Return to the Humble Point Pattern, (paper under review) for Deweirdt, E. and J. Bourgeois (eds.) *Spatial Analysis Applied to Archaeological Sites from Protohistory to the Roman Period*. Ghent.

Bevan, A. and Conolly, J. 2006. Multi-scalar Approaches to Settlement Pattern Analysis, in Lock, G. and Molyneaux B. (eds.) *Confronting Scale in Archaeology: Issues of Theory and Practice*: 217-234. New York: Springer.

Bevan, A. and Conolly, J. 2009. Modelling Spatial Heterogeneity and Nonstationarity in Artifact-Rich Landscapes, *Journal of Archaeological Science* 36.4: 956-964.

Clarke, D. (ed.) 1977. *Spatial Archaeology*. Boston: Academic Press.

Crema, E.R., Bevan, A. and Lake, M., 2010. A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record, *Journal of Archaeological Science* 37.5: 1118-1130.

Drennan, R. D. and Peterson, C. E., 2004, Comparing archaeological settlement systems with rank-size graphs: a measure of shape and statistical confidence, *Journal of Archaeological Science*, 31.5: 533-549.

Finkelstein, I. and Lederman, Z. 1997. *Highlands of Many Cultures. The Southern Samaria Survey*. Jerusalem: Graphit Press.

Finkelstein, I. and Magen, Y., 1993. *Archaeological Survey of the Hill Country of Benjamin*. Jerusalem: Israel Antiquities Authority.

Fisher, P., Wood, J. and Cheng, T. 2004. Where is Helvellyn? Fuzziness of multi-scale landscape morphology, *Transactions of the Institute of British Geographers* 29.1: 106-128.

Gibson, S., 2001. Agricultural Terraces and Settlement Expansion in the Highlands of Early Iron Age Palestine: Is There Any Correlation Between The Two? In Mazar, A. (ed.), *Studies in the Archaeology of the Iron Age in Israel and Jordan*. Sheffield: Sheffield Academic Press.

Gelfand, A. E. and Diggle, P. J. and Fuentes, M. and Guttorp, P. 2010. *Handbook of Spatial Statistics*, London: CRC/Taylor and Francis.

GRASS Development Team 2008. *Geographic Resources Analysis Support System (GRASS) Software.* Open Source Geospatial Foundation Project.
URL: http://grass.osgeo.org

Gregory, D. and J. Urry (eds.) 1985. *Social Relations and Spatial Structures*. London: Macmillan.

Habu, J., 2004. *Ancient Jomon of Japan,* Cambridge: Cambridge University Press.

Habu, J., 2008. Growth and decline in complex hunter-gatherer societies: a case study from the Jomon period Sannai Maruyama site, Japan, *Antiquity* 82: 571-584.

Hodder, I. and C. Orton 1976. *Spatial Analysis in Archaeology*, Cambridge: Cambridge University Press.

Hodder, I. 1977 Spatial studies in archaeology, *Progress in Human Geography* 1: 33-64.

Imamura, K., 1996. *Prehistoric Japan: New Perspectives on Insular East Asia* London: UCL Press.

Imamura, K., 2002. *Jomon no yutakasa to genkai.* Tokyo: Yamakawa.

Izquierdo, L. R., Izquierdo, S. S., Galán, J. M. and Santos, J. I., 2009. Techniques to Understand Computer Simulations: Markov Chain Analysis, *Journal of Artificial Societies and Social Simulations* 12.
URL: http://jasss.soc.surrey.ac.uk/12/1/6.html

Jarvis, A., Reuter, H.I. Nelson, A. and E. Guevara, 2008. *Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database*.
URL: http://srtm.csi.cgiar.org

Johnson, I., 2004. Aoristic Analysis: seeds of a new approach to mapping archaeological distributions through time, in Stadtarchäologie Wien Magistrat der Stadt Wien (ed.) *Enter the Past. The E-way into the Four Dimensions of Cultural Heritage (CAA 2003)*: 448–452. Oxford: Archaeopress.

Kobayashi, K., 2008. Jomonjidai no rekinendai. In: Kosugi, Y., Taniguchi, Y., Nishida, Y., Mizunoe, W. and Yano, K. (ed.) *Rekishi no monosashi: : Jomon jidai kenkyu no hennen taikei*, 257-269. Tokyo: Douseisha.

Lennard-Jones, J.E. 1924. On the determination of molecular fields. *Proceedings of the Royal Society of London A* 106: 463–477.

Li, X. 2012. Standardisation and Labour Organisation of the Bronze Weapons for the Qin Terracotta Warriors, China (PhD thesis, University College London)

Markofsky, S. and A. Bevan 2011. Directional analysis of surface artefact distributions. A case study from the Murghab Delta, Turkmenistan, *Journal of Archaeological Science* 39.2: 428-439.

Martinón-Torres, M., Li, X., Bevan, A. Xia, Y.. Zhao, K. and T. Rehren (in press) Forty thousand arms for a single emperor: from chemical data to the labour organization behind the bronze arrows of the Terracotta Army, *Journal of Archaeological Method and Theory*.

Mehrer, M. W. and Wescott, K. (eds.) 2006. GIS and Archaeological Predictive Modeling, London: CRC/Taylor and Francis

O'Sullivan, D. and D. Unwin 2003 *Geographic Information Analysis*, Hoboken: Wiley.

Pentedeka, A., E. Kiriatzi, L. Spencer, A. Bevan, J. Conolly 2010. From Fabrics to Island Connections: Macroscopic and Microscopic Approaches to the Prehistoric Pottery of Antikythera, *Annual of the British School at Athens* 105: 1-81.

Premo, L., 2010. Equifinality and Explanation:The Role of Agent-Based Modeling in Postpositivist Archaeology. In: Costopoulos, A. and Lake, M. (ed.) *Simulating Change: Archaeology into the Twenty-First Century*. Salt Lake City: University of Utah Press, 28-37.

R Development Core Team 2011. *R: A Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing.
URL: http://www.R-project.org/

Ratcliffe, J.H. 2000. Aoristic analysis: the spatial interpretation of unspecified temporal events, *International Journal of Geographical Information Science* 14.7: 669-679.

Robert, C. P. and Casella, G., 2004. *Monte Carlo Statistical Methods (2nd Ed.).* New York:Springer.

Vanzetti, A., Vidale, M., Gallinaro, M., Frayer, D.W. and Bondioli, L. 2010. The iceman as a burial, *Antiquity* 84: 681-692.

Verhagen, P. and Whitley, T.G. 2011 Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene, *Journal of Archaeological Method and Theory* (online first copy).

Wilson, S. M. and Melnick, D. J. 1990. Modelling Randomness in Locational Archaeology, *Journal of Archaeological Science* 17.4: 403-412

Wood, J. 2009 *LandSerf: A Geographic Information System for the Visualization and Analysis of Surfaces*, London City University
URL: http://www.landserf.org/

Zertal, A., 1988. The water factor during the Israelite Settlement Process in Canaan. In Heltzer M. and Lipinski E. (eds), *Society and Economy in the Eastern Mediterranean (c. 1500-1000 BC)*, 341-52. Leuven: Peeters.