

## Multiscalar approaches to settlement pattern analysis

Andrew Bevan and James Conolly

Postprint of a 2006 chapter in Lock, G and Molyneaux, B, (eds.) *Confronting Scale in Archaeology: Issues of Theory and Practice*, 217-34. New York: Springer.

### 1. Introduction

The increasing popularity of Geographic Information Systems/Science (GIS) in archaeology can be linked to the development of user-friendly software and corresponding improvements in spatial data handling techniques. As a result, GIS is deployed commonly as an organisational tool, but rather less attention has been paid to important developments in spatial statistics that can help make sense of such datasets. Perhaps the most important new developments all relate to the issue of scale, with respect to: (i) the problems encountered when combining datasets collected at different resolutions (e.g., Gotway and Young, 2002); (ii) the scale-related biases inherent in aggregate analytical units (e.g., the Modifiable Areal Unit Problem: Openshaw, 1996; Harris this volume); (iii) or techniques for multiscalar pattern recognition.

Here, we focus on the last of these three areas. Our paper reviews existing statistical approaches to settlement patterning in archaeology, explores in detail one useful multiscalar method – Ripley's K function – and suggests both the problems and potential of such techniques when interpreting the particular evidence provided by landscape survey. We draw on case studies from the Kythera Island Project (KIP), a multidisciplinary initiative designed to study the cultural and environmental history of the island of Kythera, Greece (Broodbank, 1999). From the project's onset in 1998, GIS has been used to store, manage, and analyse a wide variety of KIP research contributions, including the results of intensive archaeological survey, geoarchaeology, botany, historical geography and archival studies (Bevan and Conolly, 2004).

### 2. Spatial Statistics and Settlement Patterns

Settlement analysis in archaeology seeks to build up from the static spatial distribution of material culture and anthropogenic modifications visible in the contemporary landscape to an understanding of the dynamic cultural and environmental processes of human settlement systems.

With the obvious exception of phenomenological approaches, most studies of settlement and landscape accept that there is a need to adopt an empirical approach to pursuing this goal, even if in so-doing, many then fail to embed their conclusions within a wider inferential framework. Standard quantitative methods tend to explore either: (i) correlations between settlement (or other zones of human activity) and social or environmental variables (e.g., "predictive modelling"), or; (ii) the degree to which new settlements or households are located in physical relation to existing ones (we might call this "neighbourhood dependence"). The traditional tools used by archaeologists include, respectively, linear or logistic regression and nearest neighbour or quadrat analysis, but each

of these raises methodological problems. The first two have the capacity to mislead in contexts where spatial dependence can be shown to exist (i.e., most geographic contexts: Fotheringham et al., 2002:162–166), and the last two are insufficient for detecting multiscalar spatial patterns. Here we concentrate on the latter, but the need to integrate these approaches is raised again in discussion at the end.

Settlement distributions are often described in terms of their configuration vis-a-vis three idealized states – namely random, regular, or clustered – but rarely do these states occur so clearly in practice (van Andel et al., 1986). In reality, settlement patterns are more complex, and measures such of these need to be contextually sensitive to the fact that the scale of analysis can change what appears to be a nucleated or centralized pattern, to one better described as dispersed. A regular or uniform pattern between contemporaneous sites has been taken to reflect a form of competition between settlements, the existence of site catchments or both (Hodder and Orton, 1976:54–85), sometimes because of demographic growth from an initial random distribution (Perles 2000:132–147). Clustering of sites may result from a number of factors, although localized distribution of resources and the emergence of polities or regional centres have often been highlighted (Roberts, 1996:15–37; Ladefoged and Pearson 2000). In contrast, random distributions have usually been treated as the statistical null-hypothesis, though several commentators provide good examples of how apparent random distributions in fact can be conditioned by selected environmental, biological, and social variables (Maschner and Stein, 1995; Woodman 2000; Daniel 2001). Indeed, a problem we will return to later is that point pattern analysis implicitly assumes spatial isotropy (i.e., invariance by rotation) and homogeneity despite the fact that actual human landscapes offer both topographically dependent movement environments and spatially heterogeneous natural resources (water, soils, etc.).

The favoured technique of archaeologists for detecting clustered or uniform distributions is nearest neighbour analysis. Clark and Evans (1954) first explored the utility of this method for ecological purposes, and it was soon being used to understand settlement patterning (Dacey, 1960; Haggett, 1965). Its application to archaeological settlement pattern analysis followed some time later (Hodder and Hassall, 1971; Hodder, 1972; Whallon, 1973; Washburn, 1974; Hodder and Orton, 1976), continued in the 1980s and 1990s, and the technique retains its prominence today both in general textbooks (e.g., Wheatley and Gillings 2002) and culturally specific studies (e.g., Ladefoged and Pearson 2000; Perles 2001:134–138).

Clark and Evan's nearest neighbour coefficient is probably popular in the archaeological community for two reasons: (i) it is straightforward to calculate, and; (ii) it provides an easily interpretable coefficient. However, nearest neighbour analysis was not designed to detect spatial patterning at anything but the 1st nearest neighbour. Increasing the nearest neighbour measurement to the 2nd, 3rd...nth neighbour may detect clustering at different scales, but the statistical validation of patterning then becomes difficult (Hodder and Orton, 1976:41). Nearest neighbour analysis is also significantly influenced by the size

of the area to be analyzed, with regular, random, or clustered distributions arising being dependent on the amount of surrounding area included in the analysis (Hodder and Orton, 1976:41). While there are workarounds for these problems, the technique remains a relatively coarse ruler with which to measure point distribution patterns.

In particular, the focus on 1st neighbour distances may overlook more complex, multiscalar, spatial patterns. Consider, for example, the point patterns in Figure 14.1. The left panel shows a hypothetical distribution of 56 sites. A single order nearest neighbour analysis applied to the 56 points in the left panel would detect the presence of clusters, and a K-means statistic could be employed to show that the optimum number of clusters was likely to be 8 (Blankholm, 1990:65). However, neither of these analyses would be able to identify the fact that there is also a higher-order scale producing three clusters. Furthermore, if we include the finer artefact-scale resolution represented on the right panel (rather than just an approximation of the centre of the artefact distribution), then clustering can be shown to exist at three different spatial scales: (i) artefacts forming sites (clusters i-x); (ii) sites forming primary clusters (clusters 1-8); and (iii) primary clusters forming secondary clusters (clusters A-C).

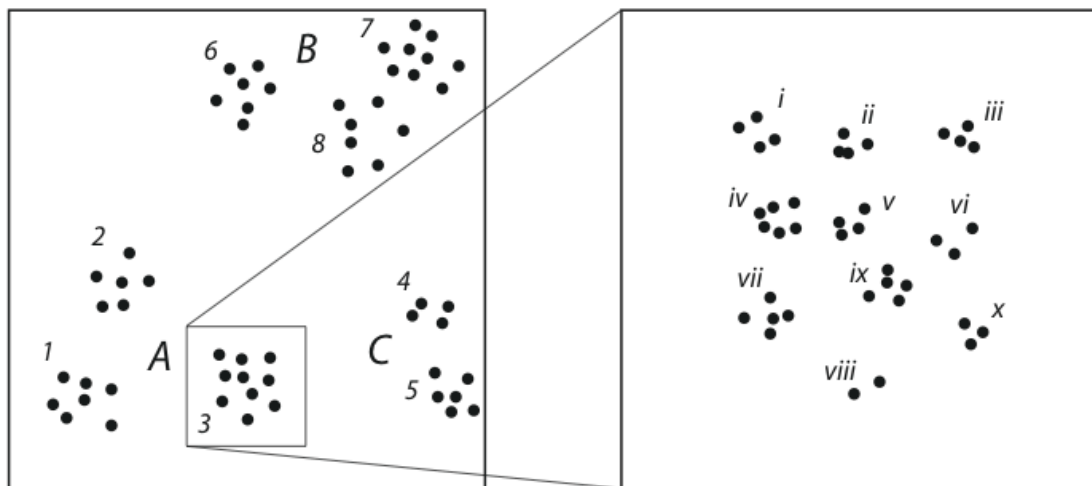


Figure 14-1. Multiscalar point patterns.

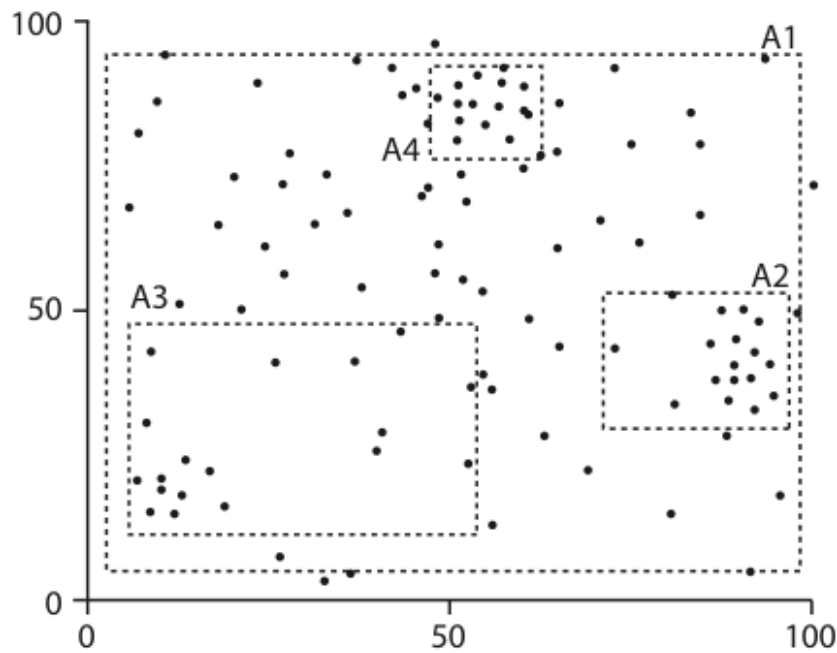


Figure 14-2. The influence of study area size on the detection and characterization of patterns (after Goreaud and Pélissier 2000: 15).

Another major problem with nearest-neighbour analysis is the effect the size of the study area has on the detection of patterning. For example, Figure 14.2 shows how adjusting the scale of analysis has a major influence on the homogeneity, intensity and clustering tendencies of point distributions. In the entire study area, A1, the pattern is homogenous with a clustered structure (i.e., clustering occurs relatively evenly) so that a frequency distribution of nearest neighbour values would be normally distributed. At smaller scales, for example in area A2, the pattern is heterogeneous with a strong left to right gradient. A neighbourhood density function would be positively skewed with a bimodal tendency. Area A3 is similarly heterogeneous, although its density value is significantly lower than A2. Area A4 has a high intensity and homogenous distribution, although here it is far more regular than seen elsewhere.

These two scalar issues – one related to analytical resolution, the other related to analytical area – although presented in abstract, are very real when attempting to make sense of settlement distributions, given that the latter may show a variety of characteristics depending on the resolution and the shape of the study area. The fact that GIS-led approaches to the collection and management of archaeological survey data are able to store data at several different scales within the same environment (e.g., artefacts, sites and regions) underlines the need for spatially sensitive approaches to the analysis of distribution patterns. Moreover, the dichotomy created by nearest neighbour analyses, dispersion vs. nucleation, is useful only at a very general level. Measures that take into account the intensity of settlement, its homogeneity, and the scale at which it is clustered or dispersed are clearly superior.

We therefore propose the use of a broader range of statistical approaches to point patterning, including methods that are inherently multiscalar such as Ripley's K-function that we consider in some detail. Ripley's K-function (Ripley, 1976; 1997; 1981) was designed to identify the relative aggregation and segregation of point data at different spatial scales. It is defined for a process of intensity  $\lambda$ , where  $\lambda K(r)$  is the expected number of neighbours in a circle of radius  $r$  at an arbitrary point in the distribution (Pélissier and Goreaud 2001:101). The K-distribution is a cumulative frequency distribution of average point intensity at set intervals of  $r$ . Significance intervals are generated by Monte Carlo simulation of random distributions of the points (Manly, 1991). These estimates can be compared with the observed values of  $K$  to provide a statistically robust measure of cluster size and cluster distance in the dataset. We use an edge effect correction method proposed by Goreaud and Pélissier (1999). For clarity of presentation, the cumulative K distribution is usually transformed to  $L(r) = \sqrt{K(r)/\pi} - r$ , where the expectation under randomness ( $L(r) = 0$ ) is a horizontal line.  $L(r) < 0$  means that there are fewer than expected neighbours at distance  $r$ , suggesting a regular pattern, and  $L(r) > 0$ , means that there are more neighbours at distance  $r$ , indicating a clustered pattern (Pélissier and Goreaud 2001:102).

### 3. Research Context

Our broad area of interest is human settlement in Mediterranean landscapes, particularly the Aegean. The Aegean was first colonized by late Pleistocene pre-modern humans, possesses the earliest farming communities in Europe and, during the Bronze Age, was the setting for some of Europe's first complex societies. It saw the rise of the Greek Classical polis-states, and was subsequently entangled in the geopolitics of the Roman, Byzantine, Ottoman, Venetian and British empires. In the 19–20th century it was brought under the umbrella of the modern Greek nation state, and most recently, in the 21st century, it is part of an emerging European super-state. The impacts of these events on Aegean rural landscape history have been the subject of an enviable breadth of intensive survey projects that have provided high-resolution data on long-term dynamics of Aegean settlement systems (e.g., Broodbank, 1999; Cherry et al., 1991; van Andel et al., 1986; Bintliff and Snodgrass, 1985; Renfrew & Wagstaff, 1982, to name but a few). The Aegean is of obvious importance for archaeologists interested in long-term patterns and processes of human social and cultural evolution, mobility and population dynamics, settlement systems and ecology.

In this study, our largest unit of analysis is the island of Kythera, which lies approximately 15 km from Cape Malea on the southern tip of the Peloponnese (Figure 14.3). Its geographical location between two distinctive and influential regions, the Greek mainland and Crete, has been instrumental in shaping a distinctive Kytherian history. The island's role within, and contribution to, wider (early Aegean, eastern and/or central Mediterranean, and later pan-European) social and economic networks was one of the several thematic issues underlying the establishment of the Kythera Island Project (KIP) in 1998 (Broodbank, 1999). One way to explore the relationship between on- and off-island processes is to consider the island's settlement patterns, particularly for cycles of nucleation and dispersion. Such cycles have been documented in other areas of

the Aegean, notably on Melos (Renfrew and Wagstaff, 1982), Kea (Cherry et al., 1991:474) and in the southern Argolid, and interpreted as strategic responses to the expansion and contraction of inter-regional trade and exchange. Although such information does not offer a simple ruler to measure the islands and islanders' relationship to (political) economic cycles, it can offer insight into how the island settlement and demographic structure responded to broader trends in Aegean social and political history.

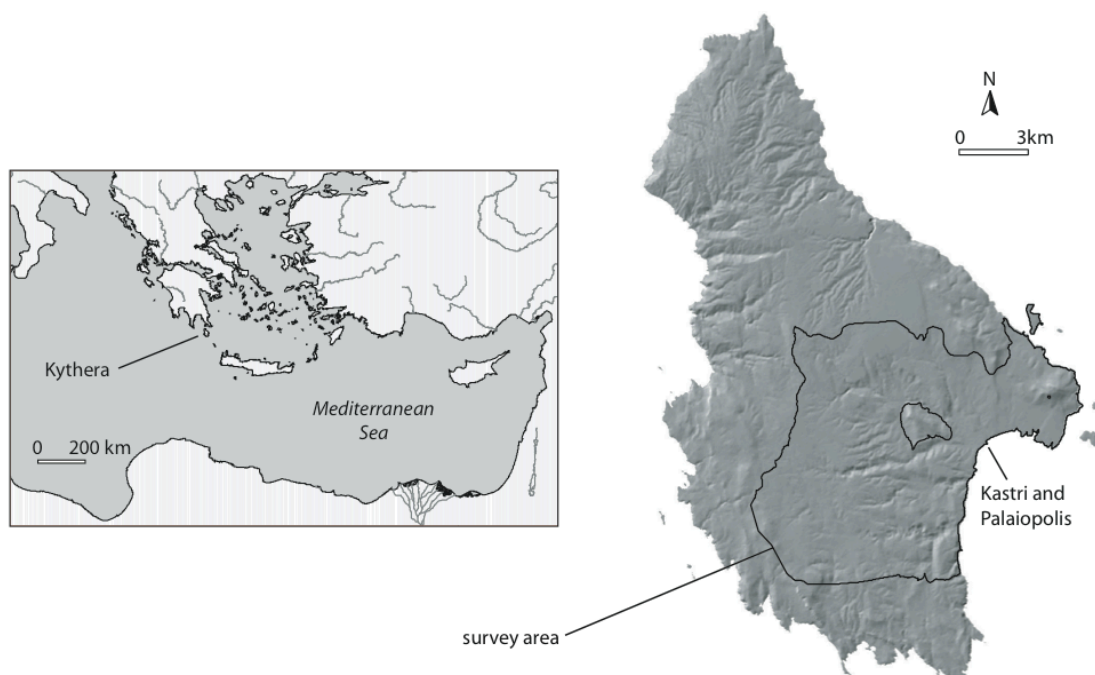


Figure 14-3. Kythera and the KIP survey area.

While our interest in exploring settlement patterning reflects this broader concern with Aegean rural landscape history, our purpose here is primarily methodological. Previous work by both of us (Bevan 2002; Bevan and Conolly 2004; Conolly 2000) have highlighted the influence of analytical scale on constructing meaning from the archaeological record and this paper offers a further contribution to this endeavour by assessing critically the statistical tools available for quantifying nucleation/dispersal phenomena. We hope that this will provide a sound analytical platform upon which further, more holistic analysis of the Kythera material may proceed, and will also be of considerable use to others who wish to make sense of such landscape patterning at different spatial scales.

#### 4. The KIP Dataset

Intensive archaeological survey between 1998 and 2001 has documented the location and chronology of nearly 200 previously unknown prehistoric, Classical, Roman, Medieval and Venetian settlements across a study area covering about one third of the island. More comprehensive information about the fieldwork can be found in Broodbank (1999) and Blackman (1999, 2000, 2001, 2002). For the purposes of this paper, we restrict ourselves to analysis of settlement distributions from four chronological periods in the island's history. The first study considers the real spatial complexities behind an obvious feature of the recent Greek landscape, nucleation of settlement into villages. The second

study then moves on to consider the additional issues raised in attempting to make sense of a settlement landscape identified purely by archaeological survey, specifically for Kythera in the Second Palace Period (“Neopalatial”, ca. 1700–1450 BC). Some additional challenges – both temporal and multivariate – facing the analysis of spatial pattern in a survey context are then introduced briefly with respect to two further chronological periods on Kythera, the Early Bronze Age (ca. 3100 –2000 BC) and the Classical (ca. 480– 323 BC). Our intention is thereby to explore the viability of point pattern analysis under conditions of increasing methodological complexity.

## 5. Results and Interpretation

### *Modern Buildings and Villages*

Our first case study draws on a relatively modern dataset, primarily based on the Greek Army’s mapping of standing buildings and villages identifiable on aerial photographs of Kythera from the 1960s (see Bevan et al., 2004). This data has been checked in the field by KIP and is relatively comprehensive (though isolated field houses were often missed). Here we examine this phenomenon in two stages, beginning with ca. 9,000 individual buildings and then considering ca. 80 larger- scale “village” clusters (Figure 14.4).

We can calculate a Clark and Evans R statistic of 0.12–0.33 (depending on whether we use a mean or median nearest neighbour value) for the spatial aggregation of individual buildings on the island, suggesting a highly clustered pattern. Indeed this is confirmed if we calculate a modified K function ( $L(r)$ ). It exhibits a large positive deviation from the upper confidence interval, even at the largest distance examined (e.g., 5 km). This indicates clustering at all scales, but more importantly, the existence of a heterogeneous pattern which is being driven by more than one type of underlying process (Figure 14.5) – indeed we might logically point to the known differences between the cultural factors influencing the spacing of buildings within villages (e.g., community values, shared resources) and those effecting the position of the more isolated fieldhouses (e.g., deliberate spacing between land holdings). This kind of result is problematic because the heavy clustering in certain areas hinders correct interpretation of smaller-scale spatial structure in others.

One way to get round this problem is to analyse village and non-village areas separately. Similarly, we can step up a typological level and consider the distribution of villages represented as single points. Such analysis will be approximate because defining which building clusters constitute “villages” is often subjective (it might be made less so by calculating a K means statistic but this may cross-cut alternative political, administrative or economic definitions of “a village”). Total estimates for the island can vary from 60 to 80 distinct communities, even in the 20th century, and here we use a relatively maximal estimate. An R-statistic of 0.74–0.84 (mean and median) suggests a slightly clustered village pattern, reflecting the fact that many settlements concentrate in inland areas next to the more suitable agricultural land. Figure 14.6 plots the complete frequency distribution of nearest neighbour distances. Expected values (the grey line) were estimated by Monte Carlo simulation (i.e. from the average of a 1,000 sets, each  $n=80$ ). Again, the observed pattern suggests that small inter-

village distances (300–400 m and 500–700 m are more frequent than we might expect from a random distribution (significant at  $p < 0.001$ , Kolmogorov–Smirnov one-sample test).

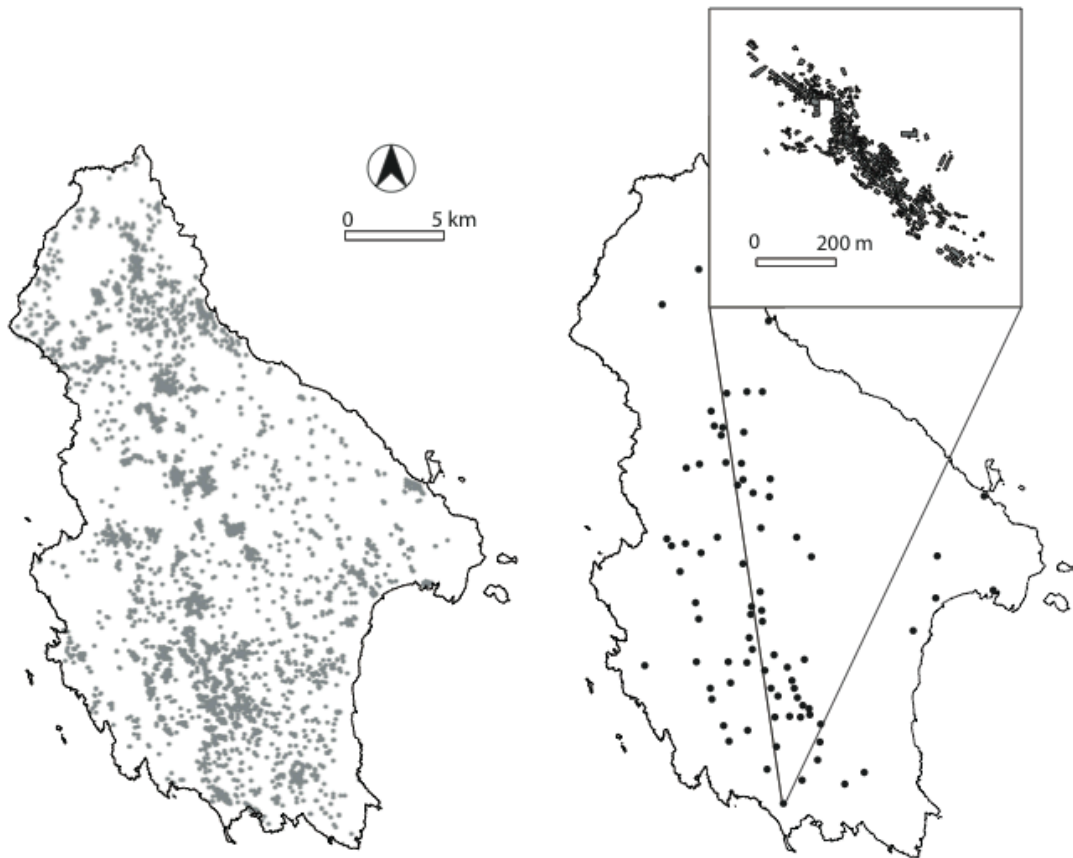


Figure 14-4. The distribution of extant buildings (left) and principal villages (right) on the island. The top right window is a close up plan of the village of Chora as an example of the original detail of the dataset.

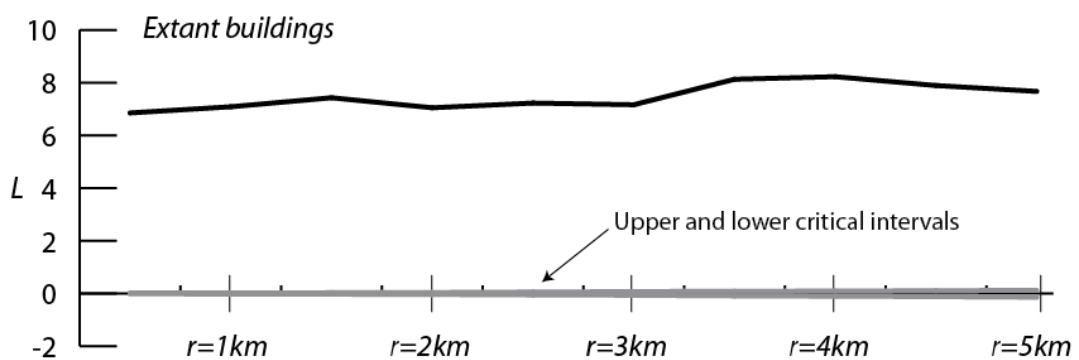


Figure 14-5. The modified K-distribution ( $L$ ) for extant buildings. The K-distribution (dark line) sits well above the (grey) line marking a clustered distribution (at  $p < 0.1$ ). Note that the upper and lower boundaries are not readily distinguishable in this chart because of the scale of the y-axis.



We can then narrow our focus to just the inland area where most villages cluster. This is useful not least because in environmental terms, this region is relatively homogenous, with similar topography, water resources and access to preferred soils. Here the spatial distribution of villages is more regular and we get an  $R=1.26$  or  $1.31$  (mean/median), for the minimum convex polygon of the inland villages. An L plot (Figure 14.7) also indicates a greater than expected regularity at smaller scales (statistically significant up to ca. 300 m radius), but in addition, shows that at larger scales, the pattern is not noticeably different from a random one. This suggests that in this more consistent resource environment, communities shared out the available space more evenly and establish clearer individual catchments, probably linked to the spatial organisation of in-field land holdings, refuse disposal and local political identity.

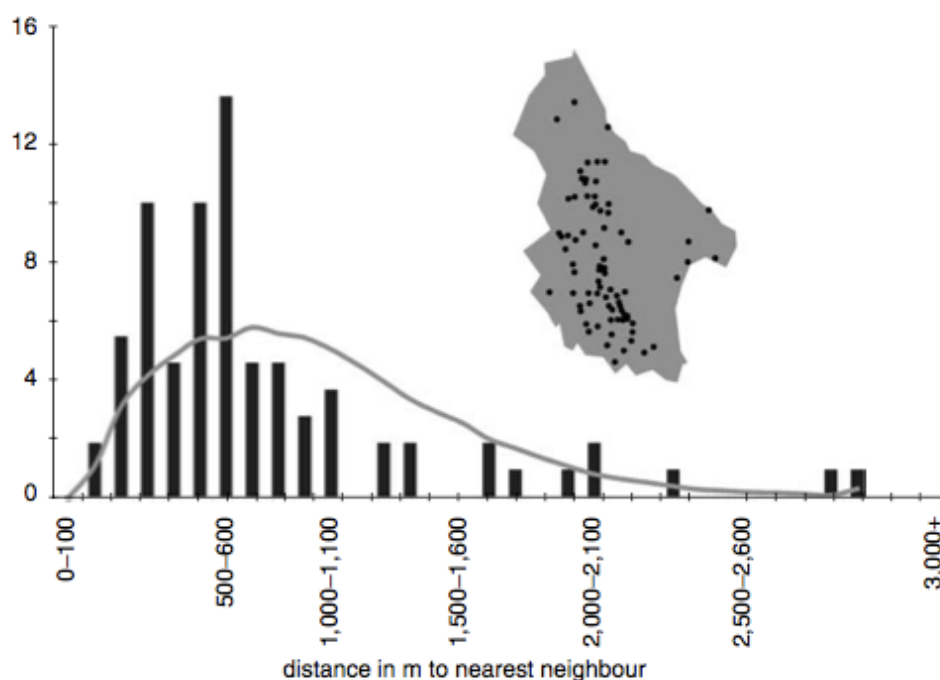


Figure 14-6. The frequency distribution of nearest neighbour distances in 100 m bins. The grey line represents an expected distribution derived from 1,000 iterations.

### *The Second Palace Period*

The previous example was chosen because it dealt with standing buildings and an extant settlement pattern (though one with some time depth). However, for most prehistoric and many historic settlement distributions, the only viable technique for accessing regional scale information of this kind is intensive surface survey. Traditional extensive reconnaissance and site recording are usually incompatible with useful point pattern analysis (except for comparing very large or prominent sites such as tells), because their coverage areas (and intensity of search) are difficult to define and because they clearly miss so much of the actual site distribution. Modern intensive survey produces far more amenable results, not only because it is concerned with expressing accurate coverage intensity, but also because, under the right conditions (geomorphological and environmental), it can hope to recover a more

comprehensive impression of past settlement landscapes. Even so, survey only produces proxy data (surface artefact scatters) for actual patterns of habitation and land use: we will therefore first explore a relatively simple, well-dated dataset—the Second Palace (“Neopalatial”, ca. 1700–1450 BC) Period sites (Figure 14.8) – before briefly considering examples in which the problems of using such proxy data are more severe. KIP has been able to document ca. 100 Second Palace sites, comprising one major port zone focused on Kastri, and beyond it, a countryside covered quite densely in small scatters (nearly 2 per km<sup>2</sup>). For a variety of reasons, these rural scatters appear to be the permanent or semi- permanent dwellings of 1–2 families engaged in agricultural subsistence (we could call them “farmsteads”: see Bevan 2002 for a preliminary analysis). They can all be quite closely dated to within a couple of centuries of each other (if not less) and therefore most are likely to have been contemporary habitations.

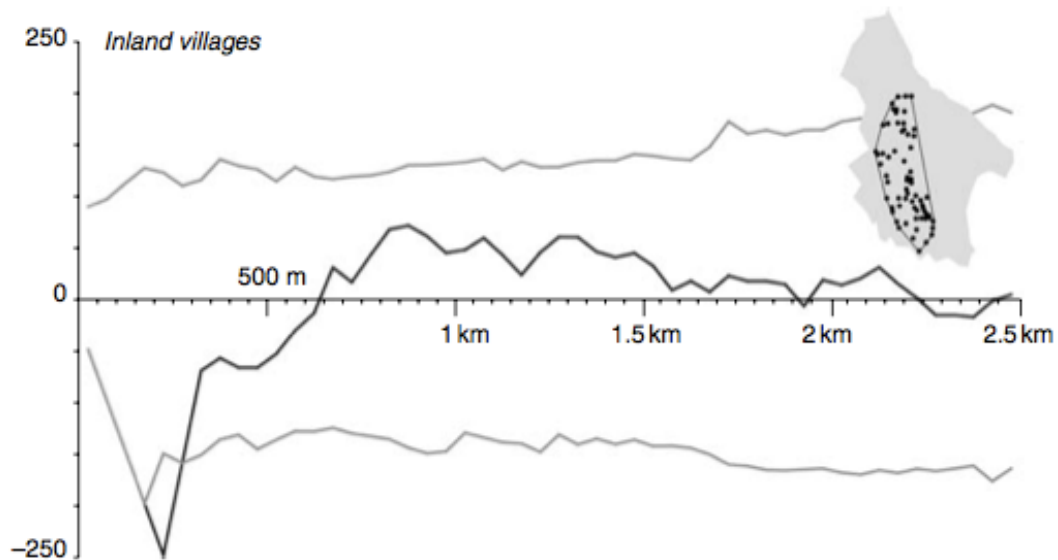


Figure 14-7. The modified K-distribution (L) for the principal villages in the inland area (as defined by the minimum convex polygon as shown). The K-distribution (dark line) sits within the upper (clustered) and lower (regular) boundaries with the exception of the range 150–300 m range (thus suggesting that at this scale there is a more regular spacing of villages than is statistically expected).

This settlement pattern is therefore a relatively simple case because it represents a comparatively shallow temporal palimpsest and comprises a limited set of site types and sizes. Even so, in order to explore the spatial distribution effectively, we must exclude the influence of the major port site of Kastri. This is because we infer that there are two separate groups of processes dictating settlement patterning in the Second Palace Period: one group generating the distribution of rural households in the hinterland of the island, and another group of processes accounting for the pattern in the vicinity of Kastri. The real influence of the clustering at Kastri becomes clear if we consider the likely distribution of actual people across the landscape as suggested by site size. If, for example, we weight the sites according to their relative size, randomly placing one point within each site scatter area for every 0.5 ha of scatter (this

threshold is arbitrary, but for our purposes here, we might think of them as crudely equivalent to notional nuclear households), then we produce a more realistic model of the likely spatial distribution of the Second Palace population. Almost all rural sites continue to be represented by a single point but the port centre (which consists of several large adjacent scatters) is represented by many more. If we run a Ripley's K function on this data, it shows a huge positive deviation from the confidence interval (Figure 14.9), again indicating significant heterogeneity.



Figure 14-8. The distribution of Second Palace Period sites within the survey area.

In contrast, if we exclude the Kastri zone as representing a separate phenomenon, we are left with a more homogenous settlement landscape, at least in terms of the theoretical population represented by each point in the analysis. We can estimate a Clark and Evans R-value of 0.85/0.91 (mean/median, where the expected values are estimated from 1,000 random sets, distributed only within the intensively surveyed area, and not the Kastri zone), which is not significantly different from a random pattern. We can also look at this across the full frequency distribution of nearest neighbour values. Previous analysis, based on the site data available in 2001 suggested certain site spacings were more common than might be expected (e.g., ca. 300 m apart: Bevan 2002:227–231). However, while these small site spacings remain highly prominent after the dataset has been augmented by further field study (adding ca. 20 extra sites), it is less clear that the pattern departs from one that might be expected by random placement. A Ripley's modified K function (Figure 14.9) suggests only a limited tendency towards regular spacing at smaller scales (significant at  $p=0.05$ , but not

at  $p=0.01$ ), and apparently random from  $r=200$  to  $500$  m. From  $r=500$  m to  $r=1,000$  m, sites cluster into statistically significant groups ( $p<0.01$ ), reflecting the exploitation of broadly preferred eco-zones, including two basins and a well-watered plateau.

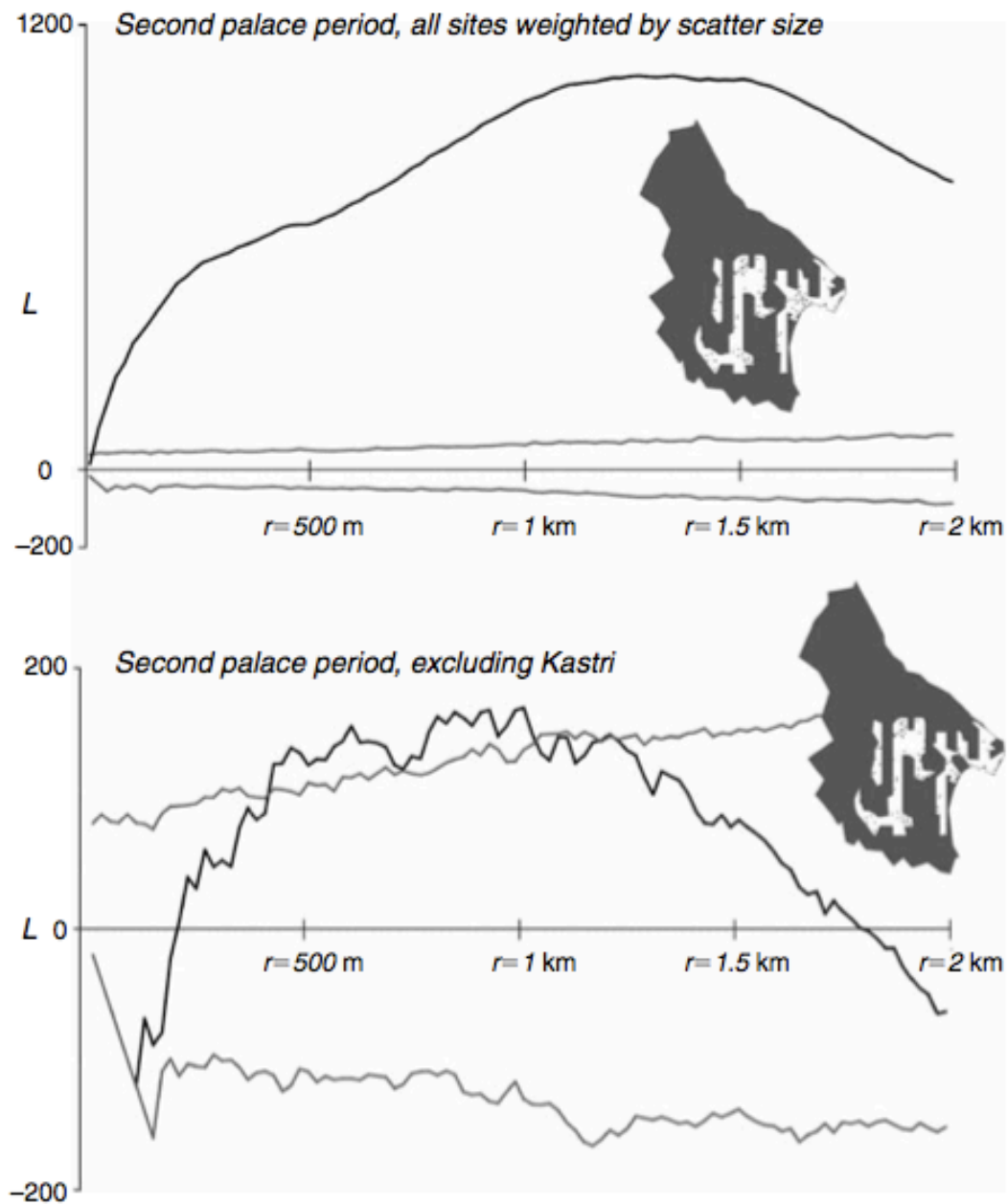


Figure 14-9. The modified K-distribution ( $L$ ) for all Second Palace Period sites (upper), and Second Palace Period sites, excluding Kastri (lower).

## 6. Prospects

Aspects of the analysis above have already suggested that a major problem in applying K-functions and related multiscale techniques to real settlement patterns is an underlying assumption of homogeneity; more precisely, such techniques assume that there is a single stochastic process behind the observed

point distribution. A typical case is the distribution created by a pure Poisson process, but there are real world examples as well (such as wind-dispersion of seedlings from parent trees or forest stand thinning: Goreaud and Pélissier 2000:27). However, in archaeological cases, we deal with a bewildering variety of heterogeneity, for example of both natural environments and of site sizes, functions and dates. Furthermore, archaeological datasets exist in anisotropic spaces (e.g., as-the-crow-flies distances rarely reflect the nature of movement across real landscapes). In contrast, point pattern analysis tends to rely on Euclidean distance measures, and while this may be an acceptable proxy at smaller scales, it becomes more problematic the greater are the distances involved.

The last of these concerns may in the future be addressed by developing more terrain sensitive ways of creating inter-point distance matrices (e.g., cost surface analysis, though not without its own problems: Douglas, 1994). This final section considers three types of heterogeneous problem in greater detail: (i) landscape variability; (ii) imperfect chronological resolution, and; (iii) variation in site size or function. Possible responses to these problems are suggested.

Landscapes vary across space in terms of their provision of water resources, soils, rainfall, solar irradiation, etc. The degree of spatial heterogeneity will itself also vary between landscapes, with highly localized heterogeneity being a defining feature of Mediterranean environments (Horden and Purcell 2000). In terms of settlement patterns, this is unfortunate because ideally, we would want to be able to distinguish static locational preferences (e.g., driven by favoured soils) from dynamic processes such as the budding of satellite settlements from parent ones. However, there are several possible solutions:

1. Deduce a study area where the influence of environment on site location is relatively homogeneous (e.g., similar slopes, soils, access to water, etc.) and only consider point patterns within this zone (for an ecological example of such an approach: Goreaud and Pélissier, 1999:435–8).
2. Define such an homogeneous study area using more formal statistical methods such as an analysis of the local density function (Fotheringham et al., 2002:138–146).
3. Weight the intensity of random points so they follow the apparent locational preferences of observed sites. This could be done by allocating random points (during Monte Carlo simulation) according to a site location probability surface (e.g., derived from predictive modelling) so that the resulting probability distribution broadly matches the observed one.

If we turn to the question of chronology, the EBA data from Kythera is a good example of a “temporally fuzzy” settlement pattern. A KIP sample of 60p sites at a density of nearly 1:5/km<sup>2</sup> represents in Aegean EBA terms a superb dataset and K function analysis suggests sites have a tendency to cluster significantly at  $r=500$  m, with clustering maintained to about  $r=1250$  m (Figure 14.10). However, it is extremely unlikely that all of these sites were in use contemporaneously or continuously throughout the ca. 1,000 year duration of

the EBA. Finer chronological resolution will be possible in the future (fabric analysis can often distinguish three EBA sub-phases: Kiriatzi 2003), but the degree of resolution will necessarily vary from site to site. A related problem is one of continuity of occupation: Whitelaw (2000:147–150) is persuasive in suggesting that in the EBA Aegean, occupation may often have been episodic rather than continuous within any given phase. Given these difficulties, EBA settlement pattern analysis is often extremely suspect, and the clustering is more likely to be a reflection of repeated occupation of favourable environments for small-scale subsistence farming, rather than any larger-scale social process.

These problems emphasize the need to find formal methods to incorporate temporal uncertainty into our pattern analysis. Not only do we have to consider distributions that have been “dated” to our period of interest with varying degrees of diagnostic certainty (both at the scale of individual artefacts and for the overall dating of sites), but we also have to contend with the possibility that individual sites may have been discontinuously occupied throughout the finest chronological divisions we can achieve and therefore may not be contemporary landscape phenomena. On a practical level, these problems benefit from a min-max approach in which analysis is run (i) on the sample of definite sites only, and (ii) on all possible sites in the phase. If there are indications that clustering/regularity persists in each of these cases then the pattern can be considered a robust one. Similarly, one response to the possibility of discontinuous settlement is to perturbate the site distribution artificially for any given chronological phase by arbitrarily excluding a certain number of points — again, if clustering or regularity persists despite repeated minor alterations of this kind (the process is a kind of internal Monte Carlo test and is necessarily laborious) then it can be interpreted with greater confidence.

The KIP Classical sites are a good instance of the problems of categorical heterogeneity. We not only have sites and settlement clusters of varying sizes, but also of varying functions, that preliminary analysis suggests probably include a major port, permanent farmsteads, temporary shelters, sanctuaries, kiln sites and metallurgical areas. We might consider the spatial relationships between any one or two of these categories, but the results (Figure 14.10) become meaningless the more functionally mixed the dataset becomes, except to show some possible heterogeneity (in that the observed pattern barely comes down within the confidence interval at large distances) and that from about  $r=250$  m there is strong clustering of (diverse) activity areas. The key is therefore to compare like with like, reducing analysis to consider only one or two categories (K-function analysis can be extended to consider bivariate spatial relationships). In methodological terms this means careful assessment of function on a site by site basis (e.g., “villas” in a Roman landscape). Likewise, even for a relatively continuous variable such as site size, physical (population), economic (permanent market) or political (formal municipal status) thresholds sometimes exist that can guide sub-classification.

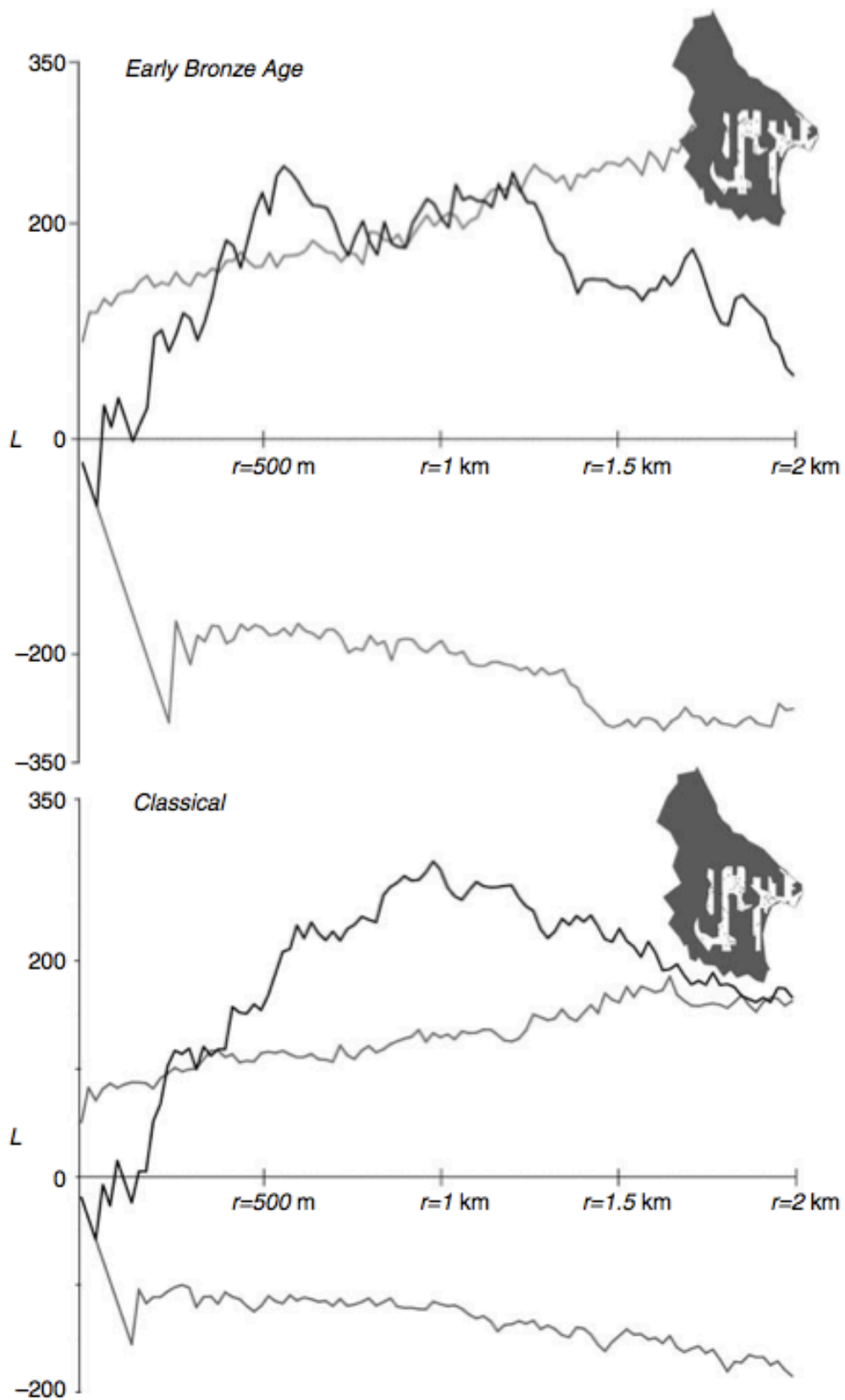


Figure 14-10. The modified K-distribution (L) for EBA sites (upper) and Classical sites (lower).

## 7. Conclusions

This paper has emphasized the highly reflexive approach necessary for the correct identification and interpretation of the processes behind settlement patterns. In our opinion, the key challenges are: (i) to define a sample/study area and its levels of search intensity appropriately (correcting for or exploring “edge effects” statistically where necessary); (ii) to assess and sub-divide site size, function and date range (analysing comparable features only and/or arbitrating uncertain cases statistically); (iii) to account for the resource structure of the landscape (either by only considering environmental homogenous sub-regions or by factoring resource preferences into the significance-testing stage of analysis), and (iv) to use techniques of analysis that are sensitive to detecting patterns at different spatial scales. The latter in particular is an area increasingly well-explored in other disciplines, but as yet with minimal impact on archaeological practice. There remains some value in Clark and Evan’s nearest neighbour function for identifying relationships between sites at one scale of analysis, but it may fail to detect larger-scale patterning. More critically, the dichotomy it encourages between “nucleated” and “dispersed” is at best an overly simplistic model and, at worst, bears little relationship to the reality of settlement organization, which at different scales can show both nucleated and dispersed components. In our Kytheran case study, there is obviously further work to be done, but even with the existing dataset, we have shown that using a combination of Monte Carlo testing, frequency distributions, local density mappings and Ripley’s K-function allows a more sensitive assessment of multiscale patterns and therefore a more critical evaluation of the processes underlying settlement distributions.

### **Acknowledgements**

Our thanks to the many people involved in KIP and in particular, to Cyprian Broodbank and Evangelia Kiriati (KIP co-directors) for help and advice at many stages. Francois Goreaud and Clive Orton also offered valuable help and advice. This research was conducted at the Institute of Archaeology, University College London, and Bevan’s contribution was made possible by a Leverhulme Trust Research Grant. This paper uses KIP data that was current in mid-2003. Site characterization and lab-based analyses are ongoing and thus our overall interpretations are provisional.

### **References**

- Bevan, A. 2002, The Rural Landscape of Neopalatial Kythera: A GIS Perspective. *Journal of Mediterranean Archaeology* 15.2:217–256.
- Bevan, A., Frederick, C., and Krahtopoulou, N., 2004, A Digital Mediterranean Countryside: GIS Approaches to the Spatial Structure of the Post-Medieval Landscape on Kythera (Greece). *Archeologia e Calcolatori* 14:217–236.
- Bevan, A., and Conolly, J., 2004, GIS, Archaeological Survey and Landscape Archaeology on the Island of Kythera, Greece. *Journal of Field Archaeology* 29:123–138.
- Bintliff, J., and Snodgrass, A., 1985, The Cambridge/Bradford Boeotia Expedition: The First Four Years. *Journal of Field Archaeology* 12:123–161.



Blackman, D., 1999, Kythera, *Archaeological Reports for 1998–1999* [British School at Athens] 45:20–21.

Blackman, D., 2000, Kythera. *Archaeological Reports for 1999–2000* [British School at Athens] 46:22–24

Blackman, D., 2001, Kythera. *Archaeological Reports for 2000–2001* [British School at Athens] 47:20–21.

Blackman, D., 2002, Kythera. *Archaeological Reports for 2001–2002* [British School at Athens] 48:16–17.

Blankholm, H.P., 1990, *Intrasite Spatial Analysis in Theory and Practice*. Aarhus, Aarhus University Press.

Broodbank, C., 1999, Kythera Survey: Preliminary Report on the 1998 Season. *Annual of the British School at Athens* 94:191–214.

Cherry, J., Davis, J.L., and Mantzourani, E., editors, 1991, *Landscape Archaeology as Long-Term History: Northern Keos in the Cycladic Islands*. Monumenta Archaeologica, Volume 16. Los Angeles.

Clark, P.J., and Evans, F.C., 1954, Distance to Nearest Neighbour as a Measure of Spatial Relationships in Populations. *Ecology* 35:444–453.

Conolly, J. 2000, Catalhöyük and the Archaeological 'Object'. In *Methodology at Catalhöyük*, edited by Ian Hodder, pp. 51–56. McDonald Institute of Archaeological Research Monographs, Cambridge.

Dacey, M.F., 1960, The Spacing of River Towns. *Annals of the Association of American Geographers* 50:59–61.

Daniel, I.R. (2001), Stone Raw Material Availability and Early Archaic Settlement in the Southeastern United States. *American Antiquity*, 66(2):237–265.

Douglas, D.H., 1994 Least Cost Path in GIS Using an Accumulated Cost Surface and Slope Lines. *Cartographica* 31(3):37–51.

Fotheringham, S.A., Brunsdon, C., and Charlton, M. 2002, *Quantitative Geography: Perspectives on Spatial Data Analysis*. Sage, London.

Goreaud, F., and Pélissier, R., 1999, On Explicit Formulas of Edge Correction for Ripley's K-Function. *Journal of Vegetation Science* 10:433–38.

Goreaud, F., and Pélissier, R. 2000, Spatial Structure Analysis of Heterogeneous Point Patterns: Examples of Application to Forest Stands. *ADS in ADE-4 Topic Documentation* 8.1:1–49.

Gotway, C.A., and Young, L.J., 2002, Combining Incompatible Spatial Data. *Journal of the American Statistical Association*. 97:632–648.

Haggett, P., 1965, *Locational Analysis in Human Geography*. E. Arnold, London.

Hodder, I., 1972, Locational Models and the Study of Romano-British Settlement. In *Models in Archaeology*, edited by D.L. Clarke, pp. 887–909. Methuen, London.

Hodder, I., and Hassall, M., 1971, The Non-Random Spacing in Romano-British Walled Towns. *Man* 6:391–407.

Hodder, I., and Orton, C., 1979, *Spatial Analysis in Archaeology*. Second Edition. Cambridge University Press, Cambridge.

Horden, P., and Purcell, N., 2000, *The Corrupting Sea: A Study of Mediterranean History*. Blackwell, Oxford.

Kiriati, E., 2003, Sherds, Fabrics and Clay Sources: Reconstructing the Ceramic Landscapes of Prehistoric Kythera, In *METRON: Measuring the Aegean Bronze Age*, edited by R. Laffineur and K. Foster, pp. 123–129. Liege, University of Liege.

Ladefoged, T., and Pearson, R., 2000, Fortified Castles on Okinawa Island during the Gusuku Period, AD1200–1600. *Antiquity* 74:404–412.

Manly, B.F.J., 1991, *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London.

Maschner, H.D.G., and Stein, J.W., 1995, Multivariate Approaches to Site Location on the Northwest Coast of North America. *Antiquity* 69(262):61–73.

Openshaw, S., 1996, Developing GIS-Relevant Zone-Based Spatial Analysis Methods. In *Spatial Analysis: Modelling in a GIS Environment*, edited by P. Longley and M. Batty, pp. 55–73. Geoinformation International, London.

Pélissier, R., and Goreaud, F., 2001, A Practical Approach to the Study of Spatial Structure In Simple Cases of Heterogeneous Vegetation. *Journal of Vegetation Science* 12:99–108.

Perles, C., 2001, *The Early Neolithic in Greece*. Cambridge University Press, Cambridge.

Renfrew, C., and Wagstaff, M., editors, 1982, *An Island Polity: The Archaeology of Exploitation in Melos*. Cambridge University Press, Cambridge.

Ripley, B.D., 1976, The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability* 13:255–266.

Ripley, B.D., 1977, Modelling Spatial Patterns. *Journal of the Royal Statistical Society* 2:172–212.

- Ripley, B.D., 1981. *Spatial Statistics*. John Wiley and Sons, New York.
- Roberts, B.K., 1996, *Landscapes of Settlement: Prehistory to the Present*. Routledge, London.
- van Andel, T.H., Runnels, C.N., and Pope, K., 1986, Five Thousand Years of Land Use and Abuse in the Southern Argolid, Greece. *Hesperia* 55:103–128.
- van Andel, T.H., and Runnels, C.N., The Evolution of Settlement in the Southern Argolid, Greece: An Economic Explanation. *Hesperia* 56 (1987), 303–334.
- Washburn, D. K., 1974, Spatial Analysis of Occupation Floors II: The Application of Nearest Neighbor Analysis. *American Antiquity* 39:16–34.
- Whallon, R., 1973, Spatial Analysis of Occupation Floors I: The Application of Nearest Neighbor Analysis of Variance. *American Antiquity* 38:266–278.
- Wheatley, D., and Gillings, M., 2002, *Spatial Technology and Archaeology: The Archaeological Applications of GIS*. Taylor & Francis, London.
- Whitelaw, T., 2000, Settlement Instability and Landscape Degradation in the Southern Aegean in the Third Millennium BC, In *Landscape and Landuse in Postglacial Greece*. Sheffield Studies in Aegean Archaeology, Volume 3, edited by P. Halstead and C. Frederick, pp. 135–161. Sheffield Academic Press, Sheffield.
- Woodman, P.E., 2000, A Predictive Model for Mesolithic Site Location on Islay using Logistic Regression and GIS. In *Hunter–Gatherer Landscape Archaeology: The Southern Hebrides Mesolithic Project 1988–98. Vol. 2: Archaeological Fieldwork on Colonsay, Computer Modelling, Experimental Archaeology, and Final Interpretations*, edited by S. Mithen, pp. 445–464. McDonald Institute for Archaeological Research, Cambridge.