16

Time and Probabilistic Reasoning in Settlement Analysis

Enrico R. Crema

Introduction

Archaeology has a long running research tradition in the study of spatial data (Hodder and Orton 1976, Rossignol and Wandsnider 1992, Bevan and Lake 2013), an endeavour that encouraged the borrowing of statistical methods from other disciplines. Many of these techniques have been adapted to answer existing questions, others fostered new lines of enquiry. Nonetheless the field of spatio-temporal analysis remains surprisingly under-developed in archaeology. Despite theoretical debates on the notion of time (e.g., Bailey 1983, Gamble 1987, Murray 1999, Holdaway and Wandsnider 2008, etc.) and a large number of studies in chronometry (e.g., Buck and Millard 2003, Lyman and O'Brein 2006), spatio-temporal statistics did not experience the same flourishing of ideas observed in spatial statistics and GIS-led analyses.

The distinction between spatial and spatio-temporal analyses is however a delusion as the former can be actually regarded as a special case where time is hold constant. This simplification is dictated by the assumption that, with other things being equal, neglecting the role of this dimension has a minimal influence on the reconstruction of the generative process behind the observed pattern. Most archaeologists would undoubtedly feel uncomfortable with such a statement. However, the great majority of regional analyses in archaeology do not formally consider the role played by time, which is often relegated as a qualitative attribute with little or no analytical function. Yet time plays a fundamental role in several ways. First, most spatial analyses start from the definition of an arbitrary slice of the time continuum, which we can refer to

Universitat Pompeu Fabra, Department of Humanities, 25-27, Ramon Trias Fargas, 08005, Bracelona.
CaSEs - Complexity and Socio-Ecological Dynamics Research Group.

UCL Institute of Archaeology, 31-34 Gordon Square, WC1H 0PY London.
  Email: e.crema@ucl.ac.uk

as *target interval*. *Events* occurring earlier are excluded from analysis, and hence their role in the emergence of the spatial pattern during the *target interval* is not considered. Second, spatial processes occurring within the *target interval* are often assumed to be stationary. Given the extensive duration of these chronological slices, this assumption is in most cases wrong, with the observed pattern being often the cumulative outcome of several distinct spatial processes. This problem is known as *time-averaging* (Stern 1994), a direct consequence of how much of the pattern we observe is a function of the way we subdivided our data (a problem comparable to the modifiable areal unit problem in geography, cf. Openshaw 1984). Third, defining the membership of each constituent unit of the spatial pattern, known as *event*, to a specific time window is not always straightforward. Temporal coordinates of an *event* can be defined by a *time-span* of possible existence delineated by a start and an end point (i.e., *terminus post quem* and *terminus ante quem*). Archaeological *events* have usually fairly large *time-spans* of existence, with low levels of accuracy and precision in their definition, and an undefined probability distribution between boundaries. As a result, the combination of arbitrary time-slicing processes and the uncertainty associated with each archaeological *event* can lead to a rather blurry picture of the past. Despite this, the general practice is to acknowledge their potential impact only outside the analytical stage.

This negligence is surprising, especially considering the large number of quantitative contributions in chronometry (see, e.g., Buck and Millard 2003, chapter 14 this volume). The lack of integration of these works in regional analyses is partly due to a research bias that promoted the usage of advanced statistical techniques to a limited number of case studies. Indeed most of these works rely almost exclusively on radiocarbon dates as their primary data (e.g., Parnell et al. 2008, Bocquet-Appel et al. 2009, Shennan et al. 2013, see also below).

However, the striking majority of archaeological data is not associated with radiocarbon or other forms of scientific dating (see discussions in Bevan et al. 2012 and Crema 2012), and are instead based on categorical definitions (e.g., "Early Bronze Age", "Late Archaic", etc.) inferred from the presence, absence, and frequencies of diagnostic cultural artefacts (Lyman and O'Brien 2006). This is particularly the case of settlement studies, where only few elements (e.g., dwellings, settlements, etc.) are coupled with direct scientific dating, a condition that limits the adoption of available quantitative methods in diachronic studies.

This chapter will first provide an overview of archaeological spatio-temporal analyses restricting its scope to their application within studies of regional settlement pattern. It will then discuss the limits imposed by temporal uncertainty and discuss a solution based on probabilistic reasoning and Monte-Carlo simulation. A case study from prehistoric Japan will then illustrate an example of this approach. The final discussion will review the limits and the potentials of spatio-temporal statistics in archaeology.

Spatio-temporal Analysis and Regional Settlement Study in Archaeology

The last decade saw a marked growth in the number of quantitative and computational studies of regional settlement pattern in archaeology. These cover a wide array of different themes including: political boundaries (Bevan 2011, Stoner 2012) and catchment areas (Ullah 2011); spatial patterning of settlement sites (Palmisano 2013);

distribution and layout of residential features within individual settlements (Crema et al. 2010, Eve and Crema 2014); settlement hierarchy (Crema 2013, Altaweel 2014); inter-settlement interaction (Evans and Rivers 2012); and settlement location (Carleton et al. 2012). These broad themes have been approached using a variety of techniques including point-pattern statistics (e.g., Bevan et al. 2013), model-based approach based on statistical physics (Bevan and Wilson 2013, Davies et al. 2014), and agent-based simulation (Kohler et al. 2007, Crema 2014).

Despite this flourishing of quantitative and computational techniques, the temporal dimension has been rarely approached. The few exceptions proposed in computational archaeology have been primarily focused on the visual representation of time, ranging from the development of stand-alone software packages (e.g., the TIMEMAP project http://www.timemap.net/) to the creation of GIS package plug-ins (Green 2011). Most mathematical and statistical studies have placed their emphasis on specific slices of the time-continuum, focusing on the identification of the generative process behind the observed pattern. This might range from the measurement of a summary statistic to a more formal approach involving the comparison of the observed pattern against precisely defined null models. The latter include simple random patterning of settlement location to complex point process models where both intensity and interaction are defined (Eve and Crema 2014). Others developed simulation-based approach where different assumptions on settlement behaviour have been mathematically defined to provide general theoretical expectations (Crema 2014) or precise and testable predictions (Evans and Rivers 2012, Bevan and Wilson 2013).

Successful examples of spatio-temporal analyses rely primarily on high quality data where temporal uncertainty is either confined to small intervals or quantified by scientific methods capable of assigning probability distributions for the timing of each *event*. One of the best-known examples of high-quality dataset is the spatial distribution of Classic Maya terminal dates, a record signalling the date of erection of the most recent monument at 47 sites dated between the 8th and 9th century AD (Kvamme 1990, Premo 2004). As the cessation of this practice is interpreted as the demise of the elite class and a marker of a profound socio-economic change, the most recent dates incised on these monuments have been used as a proxy of societal 'collapse' and their spatial distribution studied in relation to the extent of local socio-political interaction spheres and the general patterning of large-scale environmental stress. This case study fostered the application of a variety of geostatistical techniques (e.g., trend surface analysis; local and global indices of spatial autocorrelation, etc.) where 'time' was treated as a continuous variable representing the timing of the 'collapse', recorded at sample locations (i.e., archaeological sites).

The increasing availability of datasets with a large number of C14 samples have led to what has been labelled as the "third radiocarbon revolution" (Bayliss 2009:126), an extensive development of bespoke statistical techniques aimed to go beyond chronometry. The most successful example is the reconstruction of past population fluctuations by means of summed probability distributions obtained from calibrated C14 dates (e.g., Collard et al. 2010) sometimes coupled with more sophisticated analyses, such as the identification of possible causal relationship with climatic changes (e.g., using Granger causality analysis, as in Kelly et al. 2013), or an extensive usage of Monte-Carlo methods for detecting statistically significant episodes of population booms and busts, a solution which overcomes many of the problems related to sampling bias and idiosyncrasies in the calibration curve (Shennan et al. 2013, Timpson et al. 2014).

The wide availability of C14 samples has also led to a development of spatio-temporal analyses, such as spatio-temporal kernel density analysis (Grove 2011) and Bayesian hierarchical modelling (as in Onkamo et al. 2012), enabling to cover a wide variety of topics from inferences on the dynamics of reoccupation of north-western Europe at the end of the Ice Age (Blackwell and Buck 2003), to the detection of centres of renewed expansion of European Neolithic (Bocquet-Appel et al. 2009).

The great majority of archaeological data are however based on categorical definition of time that are characterised by higher levels of uncertainty in an undefined form. Dating generally involves the creation of arbitrary sets of artefact types grouped and linked to a spatio-temporal extent (i.e., an archaeological "period", "phase", or "culture") and the definition of memberships to one or more of these. Although this is a relatively straightforward exercise, ambiguities associated with the precise definition of the diagnostic traits, the state of conservation of the artefacts, or any form of subjective judgment of the membership introduces a significant degree of uncertainty (Bevan et al. 2012).

Settlement archaeology has an additional layer of uncertainty. *Events* are not directly dated, and their chronology is usually indirectly inferred from the dating of artefacts that are found in association with the target object. For example, the date of a residential feature is often derived from dates of objects recovered from its habitation floor. Thus, the dating process inherits the uncertainty of individual artefact dating, and also introduces the problem of how these objects are associated with the target *event* (Dean 1978). This process of association requires the definition of a depositional model, which specifies how objects are discarded and physically located in their archaeologically recovered contexts. This introduces further levels of uncertainty that cannot be overcome by simply improving the accuracy of artefact dating. A number of quantitative solutions have been proposed in the literature (e.g., Buck et al. 1992, Buck and Sahu 2000, Bellanger et al. 2008, Roberts et al. 2012), but the wide variety of depositional processes do not allow the development of unique universal solutions.

Settlement archaeology thus poses a highly challenging field for pursuing spatial analysis. The high quality temporal data provided by scientific dating does not provide a statistically sufficient sample size for most settlement studies, and does not solve the intricacies of indirect dating. On the other hand the relative dating used in most archaeological datasets imposes an arbitrary slicing of the time-continuum into unequally sized "phases" and "periods", with the temporal uncertainty of each *event* being rarely reported in formal fashion. These constraints in the archaeological data impede the adoption of existing spatio-temporal analysis for investigating regional settlement pattern, and highlight the importance in developing appropriate solutions.

Temporal Uncertainty, Probabilistic Reasoning, and Monte Carlo Simulation

Measuring Uncertainty

We can formalise the problems described in the previous section by defining a space-time comprising a set of *events* e, each with spatial coordinates s_1 and s_2 and a temporal coordinate t. If we ignore the spatial extent and the temporal duration of the *events*, we can characterise e as a spatio-temporal point pattern within a three-dimensional space.

However, the temporal uncertainty of most archaeological *events* does not allow us to assign a single value for *t*, which needs to be substituted by a probability distribution τ. This can be directly obtained from chemical or physical analysis (e.g., radiometric dating, thermo luminescence dating, etc.), or more frequently expressed in terms of membership to one or more archaeological periods based on a variety of methods (e.g., typological analysis, indirect dating, etc). In the latter case, τ is expressed as a vector of chronological periods P_1, P_2, ..., P_L with length *L,* each associated with a probability of membership. Thus low and high chronological uncertainty can be measured in terms of entropy; the distribution $\tau_A = \{0,0,0,1,0\}$ will have the lowest, while $\tau_B = \{0.2,0.2,0.2,0.2,0.2\}$ will have the highest level of uncertainty. This *event*-wise uncertainty can be quantified using Shannon's entropy value (Shannon 1948), as follows:

$$U_e = -\sum_{i=1}^{L} P_i \ln P_i \tag{1}$$

where U_e is bounded between 0 and $-L(L^{-1}lnL^{-1})$.

Chronological uncertainty is ideally an intrinsic and independent property of each *event*, and consequently all periods should be associated with similar levels of uncertainty. However, in many cases some chronological periods exhibit higher uncertainty than others due to intrinsic properties of the diagnostics used in the dating process. Bevan et al. (2012) explored this structural property of uncertainty in typological dating by devising two numerical indices. The first measures the overall uncertainty of each period:

$$U_j = \frac{\Sigma_{i=1}^{n} \, min \, (P_{ij}, max_{k \neq j}(P_{ik}))}{\Sigma_{i=1}^{n} P_{ij}} \tag{2}$$

where the uncertainty U_j for the period *j* is the sum of the minimum between the probability P_{ij} associated with each *event i* and the highest probability assigned to the *event* to any other period (P_{ik}), divided by the sum all probabilities of all *events* for the specific period. Low values (with a minimum of 0) of U_j suggest that *events* associated with period *j* have diagnostic independence (i.e., *events* have generally higher probabilities), while high values (with a maximum of 1) indicate the opposite. The second index measures instead the pair wise uncertainty between two periods:

$$U_{jk} = \frac{\Sigma_{i=1}^{n} \, min \, (P_{ij}, P_{ik}) \cdot 2}{P_{ij} + P_{ik}} \tag{3}$$

where U_{jk} is the ratio between the sum of overlaps between periods *j* and *k* for each event *i*, and the maximum possible overlap between the two periods. U_{jk} ranges between 0 indicating complete independence (no shared probabilities) and 1 indicating that all *events* share the same probabilities for the two periods.

Bevan and colleagues (2012) applied their indices to pottery sherds collected from an intensive survey in the Greek island of Antikythera. They showed how their data was characterised by an extremely wide range of overall uncertainty (between 0.091 to 0.961) and a pair wise uncertainty exhibiting higher values between abutting phases.

Whilst in many cases the vector of probabilities associated with each event is independent, and hence not influenced by the knowledge of any other event, there

are some situations where this condition is not met. This is the case when the *events* in question are features that exhibit a stratigraphic relationship, from which we can infer the temporal topology between *events*. This additional layer of knowledge can considerably improve the accuracy of the dating process; a line of research that has been explored by several archaeologists applying a Bayesian inferential framework (e.g., Buck et al. 1992, Ziedler et al. 1998).

Probabilistic Reasoning and Monte Carlo Simulation

Given the nature of our dataset, the application of standard spatial statistics is hindered by two issues. Firstly, if we seek to examine the spatial pattern for a specific chronological period P we need to formally integrate the uncertainty associated with each event. Secondly, comparisons between different archaeological periods need to take into account differences in their absolute duration.

One solution to the first problem is the adoption of weighted analysis (Crema et al. 2010, Grove 2011), whereby, for a given period, *events* with lower uncertainty have a lower contribution to the computation of the summary statistic. Whilst this approach allows the straightforward application of many techniques, the resulting statistics hide the underlying uncertainty of the raw data. An example can illustrate this. Consider two sets of *events*—A and B—with identical spatial coordinates but different temporal definitions. For a given period j all *events* of A have a probability of 0.95, while all *events* of B have a probability of 0.01. If we measure the mean centre of distribution of the two sets we will obtain the same spatial coordinates, though clearly the degree of confidence we would assign to the two summary statistics should be different, given the differences in their intrinsic uncertainty.

An alternative solution, that returns the uncertainty in the final summary statistic, can be achieved with the following three steps:

1) compute the probabilities π_1, π_2, ..., π_m for each of the m possible spatial configuration expected from the data in a given period P;
2) calculate the relevant summary statistic S_1, S_2, ..., S_m for each of these permutations;
3) obtain the probability of a given interval of the summary statistic $[S_\alpha, S_\beta]$ by summing the probabilities of all permutations where the $S_\alpha \leq S \leq S_\beta$ is met.

Figure 1a illustrates an example of this workflow with an artificial dataset representing the location of 12 archaeological sites, each associated with their probability of existence for a given temporal interval. The number of possible permutations m is given by L^n, where L is the number of possible states (archaeological periods) and n is the number of *events*. Since we are dealing with a single chronological period, we are interested only in two possible states (the target period P and all other periods), and hence $m = 2^{12}$. The example illustrated in Fig. 1 shows the average distance to nearest neighbour and the Clark and Evans Nearest Neighbour Index (NNI, Clark and Evans 1954) for three possible spatial configurations (Fig. 1b~d), and the probability estimate of defined intervals for both summary statistics (Fig. 1e and 1f).

The algorithm illustrated above provides the exact probability of any summary statistic, which quantitatively accounts for the combined uncertainty of all *events* within a dataset. However, m (the number of permutations) grows exponentially with the number of *events*, making the computation of the summary statistics intractable for
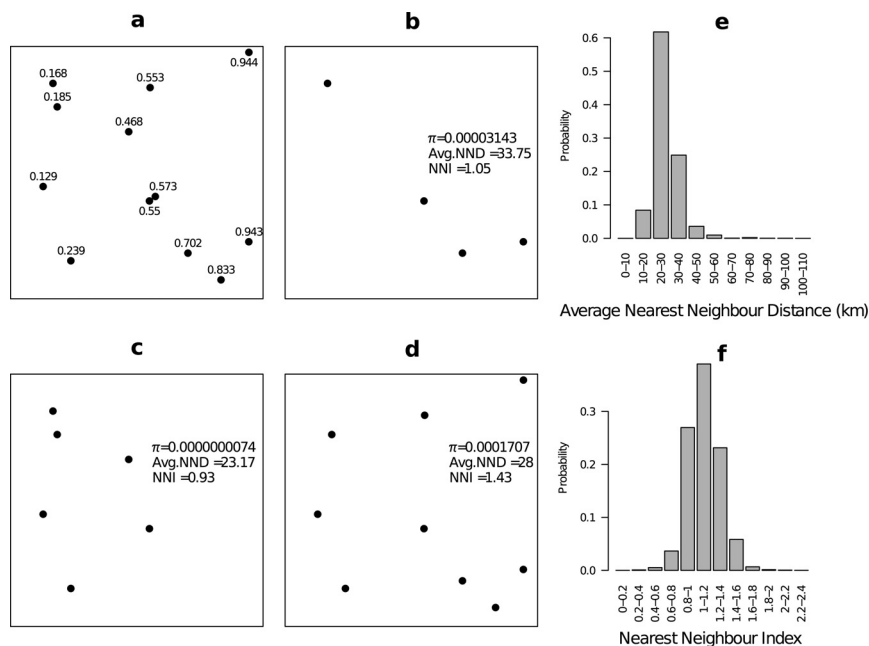
Fig. 1. An illustration of a spatial analysis with probabilistic reasoning: a) point data with probability of existence assigned to each event; b~d) three possible permutations of dataset with corresponding probability of occurrence (π), the average distance to the nearest neighbour (Avg. NND), and Clark and Evan's nearest neighbour index (NNI); e) exact probabilities for different bins of Avg. NND; and d) exact probabilities for different bins of NNI.

most datasets. One way to solve this problem is to sample from all possible permutations using the Monte-Carlo method (Crema et al. 2010, Crema 2012, Bevan et al. 2013, Yubero et al. In Press):

1) simulate k possible spatial configurations of the observed set of *events*;
2) calculate relevant summary statistics $S_1, S_2, ..., S_k$ for each of the k simulations;
3) estimate the probability of a given interval $[S_\alpha, S_\beta]$ by calculating the proportion of k satisfying the condition $S_\alpha \leq S \leq S_\beta$.

Figure 2 presents a scatter plot of the estimated probability of the condition NNI >1 using the same artificial data of Fig. 1 against different values of k (number of Monte Carlo simulations). The graph shows how, by increasing the number of simulations, the estimated probability approaches the true probability asymptotically (here shown as a dashed line). Acceptable values for k depend strongly on the strength of the pattern, the nature of the summary statistics, and the uncertainty associated with each event. Consequently, the only way to determine the optimal number of k is to plot the variation of summary statistic and determine whether the estimated probability becomes stable. Given the large values of m one might suspect that k should be similarly high, but empirical studies using this method show how this is not necessarily the case (see Crema et al. 2010).
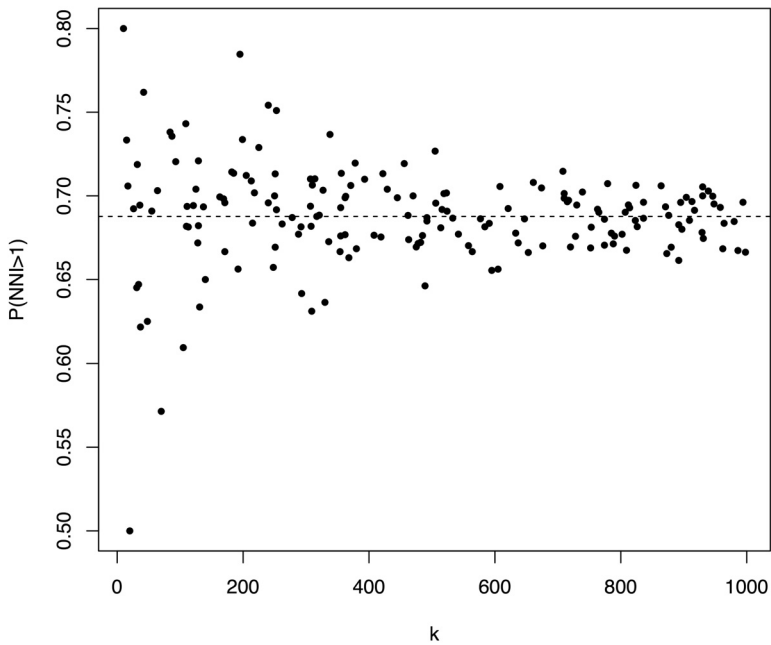
Fig. 2. Scatter-plot of *k* (number of Monte-Carlo simulations) against the probability of obtaining an NNI greater than 1. The horizontal dashed line shows the true probability.

The second issue, i.e., biases derived by comparing archaeological periods with different durations, can be solved by slicing the time-continuum into equally sized *time-blocks* (cf. with the notion of *chronon* in temporal databases, e.g., Snodgrass 1992) and extracting the probability of existence for each of them. Since these analytical units do not correspond with boundaries of archaeological periods, we need to define a continuous probability distribution of each event within their archaeological phases. This is not possible in most cases, and hence the most conservative approach is to adopt a uniform probability distribution (i.e., an equal prior) following the principle of insufficient reason (see also the application of aoristic analysis in archaeology; Johnson 2004, Crema 2012). Given that cultural traits often exhibit an initial expansion, a peak period, and a decline in popularity, adopting a "flat" probability distribution might however generate a systematic bias in the probability distribution. A unimodal shape might in fact have a closer resemblance to the true underlying probability distribution (Roberts et al. 2012, Manning et al. 2014), an assumption that plays a pivotal role in seriation techniques (Kendall 1971), and has also been integrated as a prior (using a trapezoidal distribution) in the Bayesian analysis of archaeological periods (Lee and Bronk-Ramsey 2012). However, most relative chronologies provide only the *terminus ante* and *post quem*, limiting the application of these alternative models to cases where required parameters can be estimated from the empirical record.

Case Study: Continuity and Discontinuity in the Settlement Pattern of Eastern Tokyo Bay (Japan) during the Mid-5th Millennium BP

The method described in the previous section offers a simple workflow for analysing regional settlement data using conventional spatial and spatio-temporal analyses that are widely available in the literature. Here, a case study from the Jomon culture of Japan is illustrated to showcase limits and potentials of the proposed approach. Detailed discussion on the case study, as well as a set of analyses covering a larger temporal span can be found elsewhere (Crema 2013). The choice of the case study was dictated by three compelling reasons. First, Japanese archaeology offers one of the most impressive settlement record known for prehistoric hunter-gatherers (Habu 2004), with a continuously growing dataset sustained by large scale rescue projects. These excavations provide an excellent sample for settlement studies, with multiple scales of observations ranging from the regional distribution of sites to the spatial layout of residential features within individual settlements. Second, Jomon hunter-gatherers produced large quantities of ceramic materials, which are often recovered on the floor of residential features, providing an economic (compared to radiometric methods) and effective form of indirect dating. Third, existing studies suggest that Jomon culture experienced several episodes of dramatic changes in demography and subsistence orientation (Imamura 1999), providing an ideal context where we anticipate possible changes in regional settlement pattern.

The study area is a 15×15 km square located on the Eastern shores of Tokyo Bay in Central Japan. This area is renowned for the high number of archaeological sites during the Middle (5500–4420 cal BP) and the Late (4420–3320 cal BP) Jomon periods, with the transition between the two characterised by a sharp decline in the number of residential features coupled with major changes in settlement hierarchy (Crema 2013). The present study re-examines these macro-scale trends by focusing on the continuities and discontinuities in the Jomon residential patterning during the second half of the Middle Jomon period.

The dataset comprise 934 pit houses recorded from 120 Jomon sites identified within the study area. Each residential unit has been associated with one or more archaeological phases based on published typological analysis of diagnostic pottery recovered from the habitation floor. These relative chronological assignments have then been converted into temporal intervals defined by an absolute start and end date, using the sequence proposed by Kobayashi (2004).

To reduce the potential bias generated by differently sized chronological phases, the interval between 5000 and 4400 cal BP has been sliced into 6 phases (**I**: 5000–4900; **II**: 4900–4800; **III**: 4800–4700; **IV**: 4700–4600; **V**: 4600–4500; and **VI**: 4500–4400 cal BP). Thus, the temporal definition of each residential unit has been described by a vector of eight probability estimates: 6 phases defined for this study (**I~VI**), and two dummy phases aggregating the probability of existence before (**Pre-I**) and after (**Post-VI**) our temporal scope. Probability estimates have been retrieved assuming a uniform distribution within each archaeological phase (see Fig. 3), while stratigraphic relationships between overlapping residential units have been recorded in order to ensure that the temporal topology between pit houses were conserved during the simulation process.
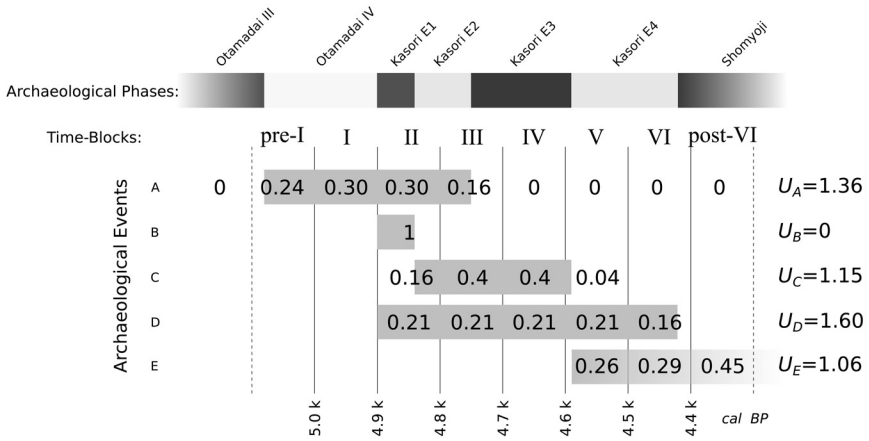
asoninganal

Fig. 3. Workflow for defining the probability of existence of each event within each phase. Each event is defined by equal membership to given archaeological periods, which has been translated into an absolute time-span of possible existence (light grey rectangles). The proportion of time-span within each time-block is then translated into probabilities. Longer time-spans have larger individual uncertainty quantified by U_e.

The distribution of event uncertainty (Fig. 4a) shows a strong peak in intermediate values, with less than 10% of data having $U_e = 0$, and no cases where the entropy value is the highest (shown here as a dashed vertical line). The overall uncertainties of the six phases show in most cases intermediate values (Fig. 4b), with the exception of phase **II** (4900–4800 cal BP), which exhibits a peak value of 0.9. This suggests that probability estimates attributed to *events* in this period are in most cases lower than other dominant phases (i.e., pit houses rarely have their highest probability of existence during this phase). The analysis of the pair wise uncertainty (Table 1) confirms the high uncertainty associated with phase **II**, with a high value of U_{jk} (=0.75) with the previous phase **I**. On the other hand, Table 1 also indicates that most of the uncertainty is distributed between immediately adjacent phases, indicating that this is rarely shared across more than two consecutive phases.

Given the large number of possible permutations ($m = 8^{934}$), computing the probability of each possible spatial configuration is not a viable option, and hence the Monte Carlo simulation approach has been adopted. Convergence test indicated that 1,000 iterations are sufficient to obtain a stable estimate of the probabilities of relevant summary statistics.

Previous analyses (see Crema 2013) on a larger dataset have indicated two key findings. First, settlement hierarchy exhibits a strong fluctuation, with an alternation between phases with few large sites and many smaller settlements, and episodes of a more uniform distribution of sizes. Figure 5 depicts a standardised rank-size plot for all 1,000 iterations with the hypothetical Zipfian (i.e., settlements of rank *r* have size Sr^{-1}, with *S* being the size of the largest settlement; Zipf 1949) drawn as a dashed line. During phase **I**, the observed rank-size distribution has a steeper line compared to the one expected by the Zipfian distribution. In the subsequent phases the size distribution becomes increasingly less hierarchical until phase **IV**, when the opposite trend starts. This fluctuation can be measured using the *A*-coefficient (Drennan and Peterson 2004,
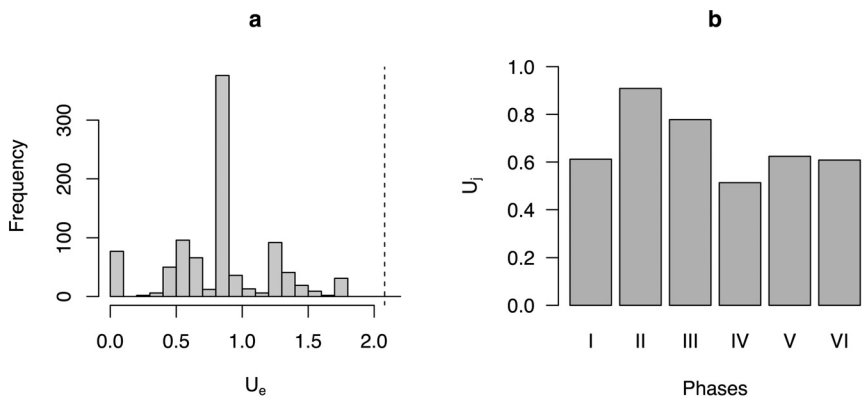
Fig. 4. Distribution of event uncertainty (a) and the overall uncertainty of each phase (b).

Table 1. Matrix of pair wise uncertainty.

	Pre-I	I	II	III	IV	VI	Post-VI
I	0.29						
II	0.15	0.37					
III	0.07	0.22	0.75				
IV	0.04	0.1	0.18	0.22			
V	0.03	0.08	0.08	0.08	0.39		
VI	0.01	0.03	0.03	0.03	0.09	0.4	
Post-VI	0.33	0.03	0.03	0.03	0.09	0.19	0.29

Crema 2013), a summary statistic that measures the amount of deviation from the Zipfian distribution. The index returns negative values for strong hierarchy, values close to zero for patterns conforming to Zipf's law, and positive values for more uniform distributions. Despite the effects of temporal uncertainty, the distribution of *A*-coefficients (Fig. 5) shows clear Gaussian shape and small variances for all periods.

Second, the total number of residential units shows an initial growth, followed by a sharp decline between phases **V** (4600–4500 cal BP) and **VI** (4500–4400 cal BP), with an average rate of decline at 0.68 pit houses per year. Figure 6 shows the box-plot of the five transitions of interests, with whiskers indicating the minimum and the maximum rates of change among the 1,000 simulation runs. The first four transitions all show positive median values, albeit high levels of uncertainty can be observed in the transition from phase **IV** (4700–4600 cal BP) to **V** (4600–4500 cal BP). The last transition suggests instead a sharp decline in the number of residential units, with a complete absence of simulation runs with positive rates of change.

These results indicate continuous changes in the settlement pattern, whereby the steady growth in residential density from phase **I** to **V** is coupled with shifts in hierarchy. However, these global statistics beg the question as to whether the variation in the settlement hierarchy was the result of a differential growth rate between different settlements or the outcome of a sharp discontinuity characterised by the formation of new settlements and the demise of others. More generally, these global statistics do not show whether temporal variations in residential density exhibit any form of
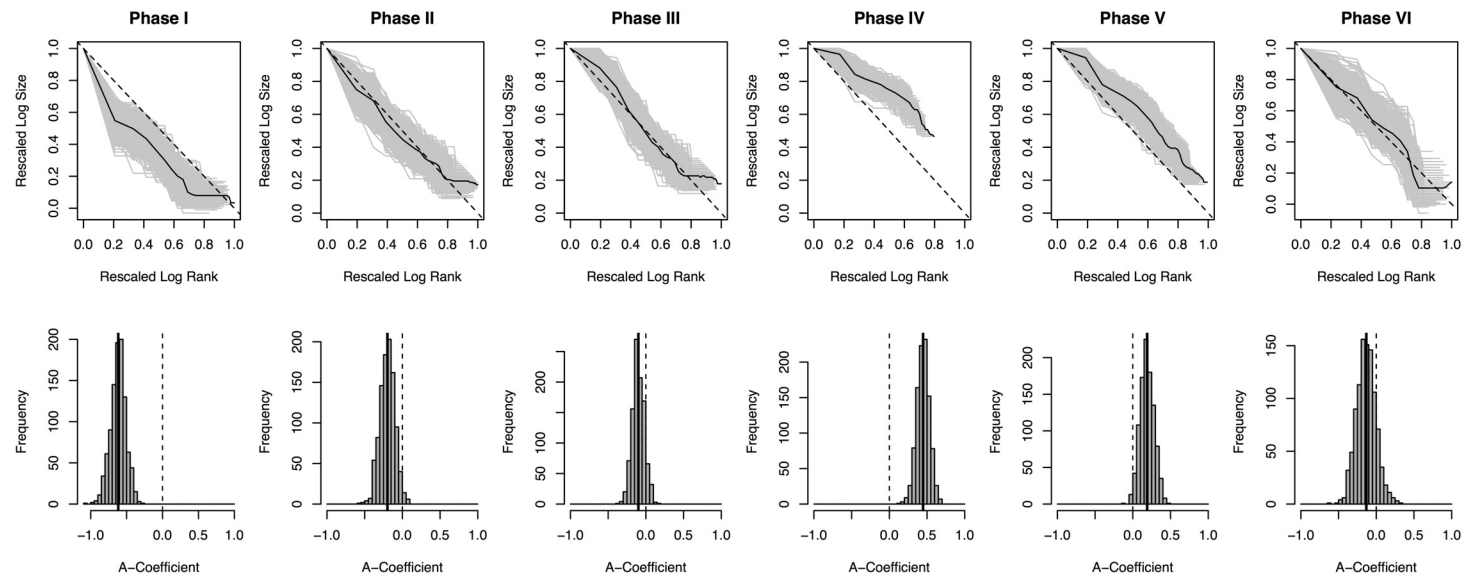
Fig. 5. Upper row: standardised rank size plot of the settlement sizes. Light-grey lines indicate individual runs of the simulation and the solid black line represents the mean trend. Lower row: distribution of the *A*-coefficient.
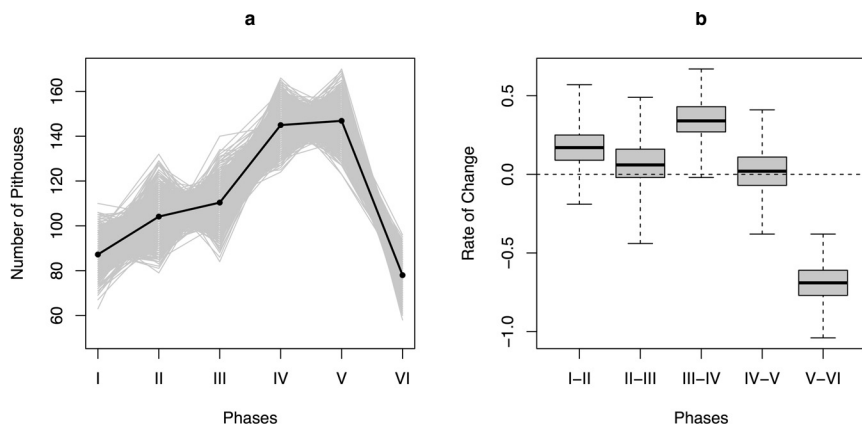
Fig. 6. Variation in the overall number of residential units (a) and corresponding rates of change (b).

spatial structure. If the fluctuations in the rank-size curves are the result of a mixture of residential growth and decline, are these opposite trends spatially clustered? If so, are they the result of local changes in the environment inducing the abandonment or colonisation of specific areas?

The temporal continuity of settlement occupation can be assessed using a variant of the Spatio-Temporal Join-Count Statistic (*st-JCS*), originally developed by Little and Dale (1999). The rational of the method can be summarised as follows. For a given transition from *i* to *j*, we first compute J_{ij}, the number of locations that are occupied by at least one residential unit in both periods. We then permute the occupied locations of both periods for *n* times, and generate a distribution of expected number of joins $\hat{J}_{ij}$ given a null hypothesis of random settlement locations. Statistical significance for $J_{ij} > \hat{J}_{ij}$ will provide a simple assessment of the continuity in the choice of residential sites. Figure 7 shows the distribution of the p-values obtained from 1,000 permutations test conducted for each of 1,000 simulated spatio-temporal patterns. The results strongly support the presence of continuity in occupation for all transitions except for the one between phase **III** and **IV**, where less than 15% of the runs had a significance level below 0.05. Interestingly, the lack of continuity[1] in the this key transition matches the highest change in the settlement hierarchy as seen in Fig. 5, when the mean *A*-coefficient shifts from slightly negative values (indicating weak hierarchy) to positive values (suggesting a more uniform distribution of settlement sizes; Fig. 5), and the period of highest growth in the overall residential density (Fig. 6).

As mentioned earlier, the global analysis depicted in Fig. 6 obscures the presence of any spatial variation in the rates of change of residential density. Figure 8 illustrates this problem by plotting the frequency distribution of rates of change between consecutive phases for each location, excluding all cases where sites were unoccupied in both phases. The histograms show how positive and negative rates of change co-exist in all transitions, indicating how the box plot in Fig. 6 hides a mixture of opposite trajectories in the change of residential density.

---

[1] Notice that the transition however does not exhibit any evidence of significant discontinuity (i.e., $J_{ij} < \hat{J}_{ij}$).
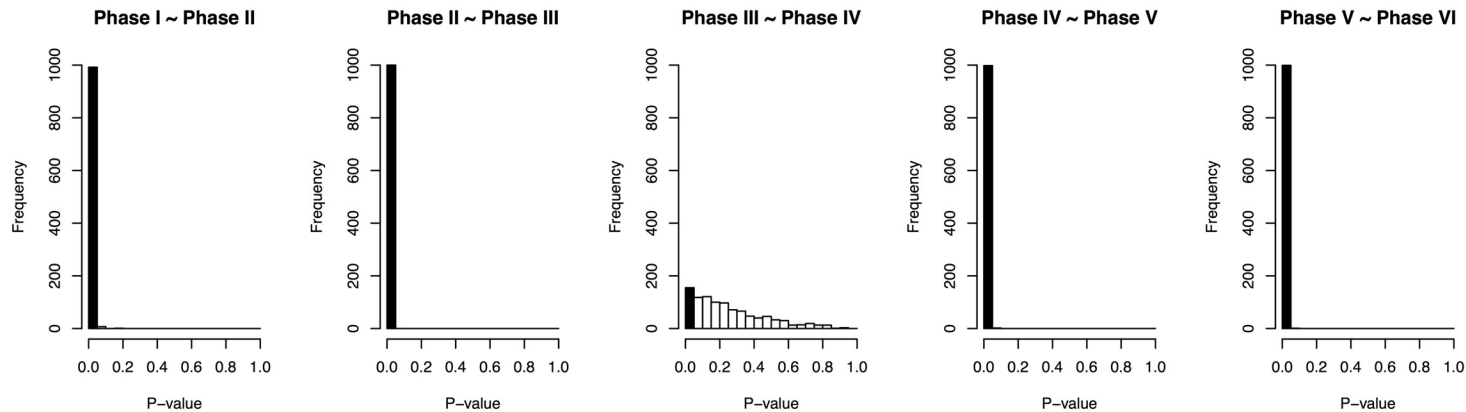
Fig. 7. Distribution of the p-values for the spatio-temporal Join Count Statistic (*st-JCS*).
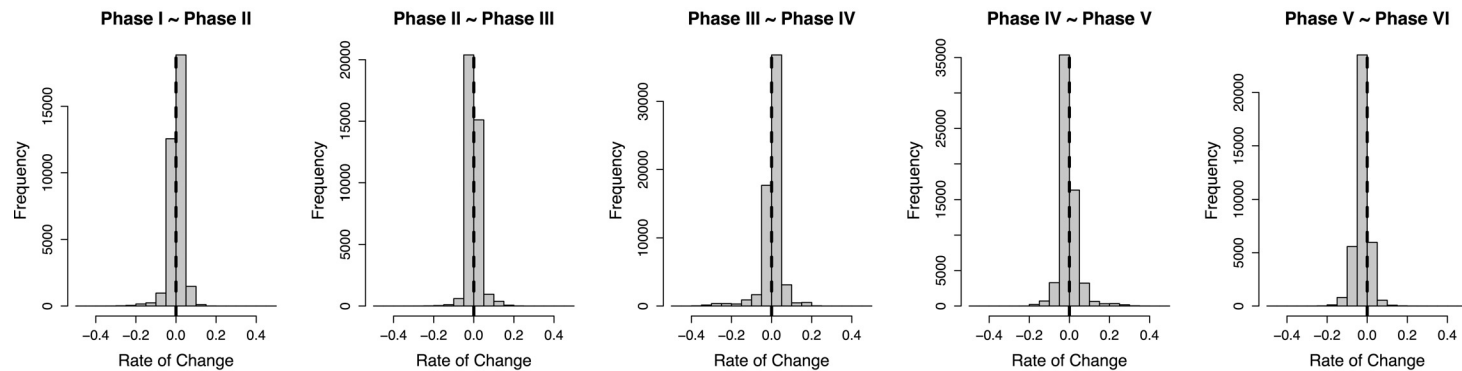
Mathematics and Archaeology

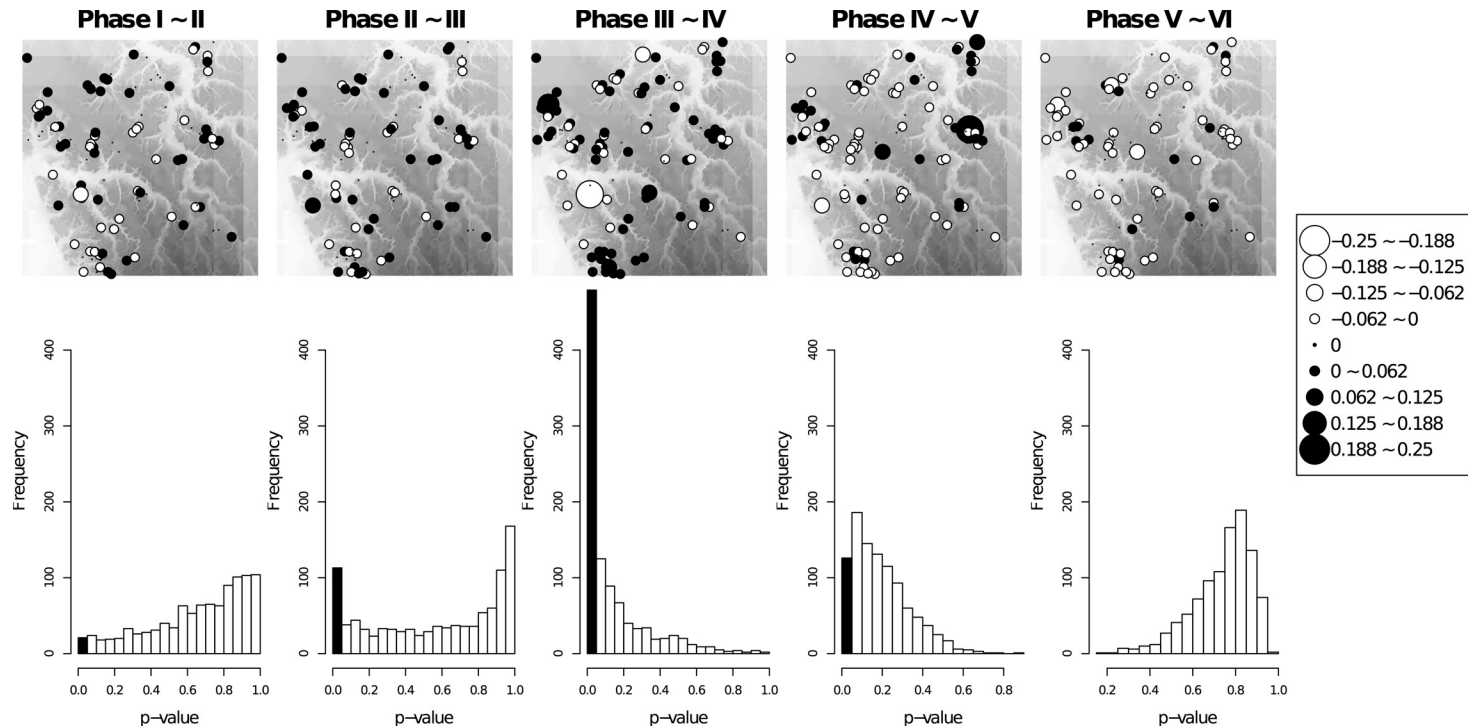**Fig. 8.** Distribution of rates of site level changes in residential density.

Fig. 9. Upper row: spatial variation in the rates of change in residential density with point size indicating magnitude and colour indicating direction (white: decrease; black: increase). Lower row: distribution of p-values for Moran's I spatial autocorrelation test. Frequencies of runs with p<0.05 are highlighted in black.

The spatial heterogeneity in the temporal variation of residential density can also be visualised as bubble plots (Fig. 9), with point size indicating the mean magnitude of change, and the colour indicating whether the trend is positive (black) or negative (white). All transitions displayed on Fig. 9 show a mixture of locations with opposite trajectories of changes in residential density, suggesting how the overall variation in the number of pit houses was characterised by a complex spatial structure. We can formally examine this by evaluating whether the rates of change in residential density exhibit spatial autocorrelation. Statistically significant results would suggest that local differences in the underlying environment might be a potential driver in the evolution of settlement pattern. The histogram on the lower row in Fig. 9 shows the distribution of p-values of Moran's I spatial autocorrelation test (Moran 1950), with a distance radius set at 2.5 km. The results suggest how the highest number of significant positive autocorrelation is observed during the transition from phase **III** and **IV**, when ca. 50% of the simulations had a p-value below 0.05. This corresponds to the highest overall growth rate in the number of residential units (Fig. 6), the largest change in settlement hierarchy (Fig. 5), and the only transition where *st-JCS* did not exhibit significant continuity in site occupation (Fig. 7). These results suggest that this period was characterised by a major change in the settlement patterning, with the concurrent abandonment of key areas and the colonisation of others, perhaps driven by episodes of local resource depletion (Toizumi and Nishino 1999). The transition from stage **V** and **VI** is instead characterised by the smallest number of runs with significant p-values. The location of sites with positive rates of change seems to be mixed with those showing the opposite pattern rather than being confined in specific areas.

Discussion and Conclusion

The analyses presented in the case study highlight both the benefits and the limitations of the Monte-Carlo approach proposed here. Despite the presence of relatively high levels of temporal uncertainty in the archaeological data and number of possible permutations, a sample of 1,000 simulation runs was sufficient to generate robust outputs in the summary statistics that can help understand the settlement evolution in the case study area. The method enabled diachronic comparisons using an array of statistics that have been already used in archaeology (e.g., Spatial autocorrelation, Rank-Size Analysis, *A*-coefficient) as well as the introduction of methods that has not been applied before (e.g., spatio-temporal Join Count Statistic).

Yet these techniques also highlighted intrinsic problems that are present in this and other datasets. The temporal slicing and the definition of arbitrary temporal blocks of 100 years have undoubtedly facilitated the comparison between different settlement configurations, but at the same time showed the intrinsic limits of the archaeological record for answering specific research questions. For example, during the transition from phase **V** to **VI**, a lack of spatial autocorrelation and the co-presence of sites with opposite rates of change in residential density have been noticed. This does not necessarily indicate that these opposite trends were concurrent, as one process might have occurred earlier than the other. In other words, the available data does not provide enough information to determine whether the spatial process within the two time-blocks was stationary or non-stationary, hindering part of our understanding of the settlement

history of the region. This is not an intrinsic problem of the proposed method, which instead highlights the limits of the available data.

In fact, the dataset used in the present case study highlights other, deeper problems in the archaeological record that limit the straightforward adoption of probabilistic technique proposed here. For example, the precise definition of the chronological boundaries of the archaeological phases adopted in the case study (based on Kobayashi 2004) hides the underlying uncertainty associated with these chronologies. Archaeologists are aware of the fuzzy nature of these boundaries and that uniform distribution does not correctly represent the assumption of a unimodal shape in the probability of existence of *events* within an archaeological phase. However, these crucial data are rarely available, and when they are present, they are retrieved only within narrow contexts where advanced statistical techniques can be coupled with robust chronological proxies. In the present case study, this enforced the use of a uniform probability distribution, which, at the current state of archaeological knowledge, represents our best approximation.

The lack of interest in defining and reporting uncertainties in chronometric data is perhaps the key reason why the development of spatio-temporal statistic in archaeology is still in its infancy. There is a wide misconception amongst archaeologists that uncertainty is a negative aspect of their research, a weakness of our record that should be kept hidden when possible. This attitude can lead to a substantial loss of information, and a reduced effort in attempting to describe, quantify, and ultimately overcome the problem of uncertainty. Crisp boundaries with no explicit definition of the underlying probability distribution are the preferred chronological description of archaeological phases, and quantitative approaches to the problem are seen as an unnecessary complication with no practical use (but see Manning et al. 2014). This trend is perhaps due to the lack of methods that can benefit from the quantification of archaeological uncertainties. This chapter is an attempt to illustrate that archaeological research can highly benefit from a stronger awareness of uncertainty, and that a development of this research agenda is highly desirable.

Acknowledgments

References Cited

Altaweel, Mark. (In press) Settlement Dynamics and Hierarchy from Agent Decision-Making: A Method Derived from Entropy Maximization. J Arch. Method. doi:10.1007/10816-014-9219-6.

TIMEMAP project http://www.timemap.net/.

Bailey, G.N. 1983. Concepts of Time in Quaternary Prehistory. Annu. Rev. Anthropol. 12: 165–192.

Bayliss, A. 2009. Rolling out revolution: using radiocarbon dating in archaeology. Radiocarbon 51: 123–147.

Bellanger, L., R. Tomassone and P. Husi. 2008. A Statistical Approach for Dating Archaeological Contexts, J. Data. Sci. 6: 135–154.

Bevan, A. 2011. Computational models for understanding movement and territory. pp. 383–394. *In*: V. Mayoral Herrera and S. Celestino Pérez (eds.). Tecnologías de Información Geográfica y Análisis Arqueológico del Territorio, Anejos de Archivo Español de Arqueología, Mérida.

Bevan, A., J. Conolly, C. Hennig, A. Johnston, A. Quercia, L. Spencer and J. Vroom. 2012. Measuring Chronological Uncertainty in Intensive Survey Finds, Archaeometry 55: 318–328.

Bevan, A., E. Crema, X. Li and A. Palmisano. 2013. Intensities, Interactions and Uncertainties: Some New Approaches to Archaeological Distributions. pp. 27–52. *In*: A. Bevan and M. Lake (eds.). Computational Approaches to Archaeological Space, Left Coast Press, Walnut Creek.

Bevan, A. and M. Lake (eds.). 2013. Computational Approaches to Archaeological Space, Left Coast Press, Walnut Creek.

Bevan, A. and A. Wilson. 2013. Models of Settlement Hierarchy Based on Partial Evidence. J. Archaeol. Sci. 40: 2415–2427.

Blackwell, P.G. and C.E. Buck. 2003. The Late Glacial human reoccupation of north-western Europe: new approaches to space-time modelling, Antiquity 77: 232–240.

Bocquet-Appel, J.-P., S. Naji, M.V. Linden and J.K. Kozlowski. 2009. Detection of diffusion and contact zones of early farming in Europe from the space-time distribution of 14C dates. J. Archaeol. Sci. 36: 807–820.

Buck, C.E., C.D. Litton and A.F.M. Smith. 1992. Calibration of Radiocarbon Results Pertaining to Related Archaeological *Events*, J. Archaeol. Sci. 19: 497–512.

Buck, C. and A. Millard (eds.). 2003. Tools for Constructing Chronologies: Crossing Disciplinary Boundaries, Springer Verlag, London.

Buck, C. and S.K. Sahu. 2000. Bayesian models of relative archaeological chronology building. Appl. Stat. 49: 423–440.

Carleton, W.C., J. Conolly and G. Iannone. 2012. A locally-adaptive model of archaeological potential (LAMAP), J. Archaeol. Sci. 39: 3371–3385.

Collard, M., K. Edinborough, S. Shennan and M.G. Thomas. 2010. Radiocarbon evidence indicates that migrants introduced farming to Britain, J. Archaeol. Sci. 37: 866–870.

Clark, P. and F. Evans. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology 35: 445–453.

Crema, E.R., A. Bevan and M. Lake. 2010. A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record, J. Archaeol. Sci. 37: 1118–1130.

Crema, E.R. 2012. Modelling Temporal Uncertainty in Archaeological Analysis, J. Archaeol. Method. Th. 19: 440–461.

Crema, E.R. 2013. Cycles of change in Jomon settlement: a case study from Eastern Tokyo Bay, Antiquity 87: 1169–1181.

Crema, E.R. 2014. A simulation model of fission-fusion dynamics and long-term settlement change. J. Archaeol. Method. Th. 21: 385–404.

Davies, T., A. Wilson, A. Palmisano, M. Altaweel and K. Radner. 2014. Application of an entropy maximizing and dynamics model for understanding settlement structure: the Khabur Triangle in the Middle Bronze and Iron Ages, J. Archaeol. Sci. 43: 141–154.

Dean, J.S. 1978. Independent Dating in Archaeological Analysis. Adv. Archaeol. Method. Th. 1: 223–255.

Drennan, R.D. and C.E. Peterson. 2004. Comparing archaeological settlement systems with rank-size graphs: a measure of shape and statistical confidence, J. Archaeol. Sci. 31: 533–549.

Evans, T.S. and R.J. Rivers. 2012. Interactions in Space For Archaeological Models, Adv. Complex Syst. 15: 1150009.

Eve, S. and E.R. Crema. 2014. A house with a view? Multi-model inference, visibility fields, and point process analysis of a Bronze Age settlement on Leskernick Hill (Cornwall, UK), J. Archaeol. Sci. 43: 267–277.

Gamble, C. 1987. Archaeology, geography and time, Prog. Hum. Geogr. 11: 227–246.

Green, C.T. 2011. Winding Dali's clock: the construction of a fuzzy temporal-GIS for archaeology Archaeopress, Oxford.

Grove, M. 2011. A Spatio-Temporal Kernel Method for Mapping Change in Prehistoric Land-Use Patterns, Archaeometry 53: 1012–1030.

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

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

Holdaway, S. and L. Wandsnider (eds.). 2008. Time in archaeology: time perspectivism revisited, University of Utah Press, Salt Lake City.

Imamura, K. 1999. Jomon no jitsuzo wo motomete, Yoshikawakoubunkan, Tokyo (In Japanese).

Johnson, I. 2004. Aoristic Analysis: seeds of a new approach to mapping archaeological distributions through time. pp. 448–452. *In*: F. Ausserer, W. Börner, M. Goriany and L. Karlhuber-Vöckl

(eds.). [Enter the Past] the E-way into the Four Dimensions of Cultural Heritage: CAA2003. BAR International Series 1227. Archaeopress, Oxford.

Kelly, R.L., T.A. Surovell, B.N. Shuman and G.M. Smith. 2013. A continuous climatic impact on Holocene human population in the Rocky Mountains, Proc. Natl. Acad. Sci. USA 110: 443–447.

Kendall. D.G. 1971. Seriation from abundance matrices. pp. 215–252. *In*: F.R. Hodson, D.G. Kendall and P. Tautu (eds.). Mathematics in the Archaeological and Historical Sciences, Edinburgh, Edinburgh University Press.

Kobayashi, K. 2004. Jomonkenkyu no shinsiten: tanso14nenndaiokutei no riyo, Rokuichishobo, Tokyo (In Japanese).

Kohler, T.A., C.D. Johnson, M. Varien, S. Ottman, R. Reynolds, Z. Kobti, J. Cowan, K. Kolm, S. Smith and L. Yap. 2007. Settlement Ecodynamics in the Prehispanic Central Mesa Verde Region. pp. 61–104. *In*: T.A. Kohler and S.E.v.d. Leeuw (eds.). The Model-based Archaeology of Socionatural Systems, SAR Press, Sante Fe.

Kvamme, K.L. 1990. Spatial autocorrelation and the Classic Maya collapse revisited, J. Archaeol. Sci. 17: 197–207.

Lee, S. and C. Bronk-Ramsey. 2012. Development and applications of the trapezoidal model for archaeological chronologies. Radiocarbon 54: 107–122.

Little, L.R. and M.R.T. Dale. 1999. A method for analysing spatio-temporal pattern in plant establishment, tested on a Populus balsamifera clone, J. Ecol. 87: 620–627.

Lyman, R.L. and M.J. O'Brien. 2006. Measuring Time with Artefacts: A History of Methods in American Archaeology, University of Nebraska Press, Lincoln.

Manning, K., A. Timpson, S. College, E.R. Crema, K. Edinborough, T. Kerig and S. Shennan. In Press. The chronology of culture. A comparative assessment of European Neolithic dating methods, Antiquity.

Moran, P.A.P. 1950. Notes on Continuous Stochastic Phenomena. Biometrika 37: 17–23.

Murray, T. 1999. Time and Archaeology, Routledge, London.

Onkamo, P., J. Kammonen, P. Pesonen, T. Sundell, E. Moltchanova, M. Oinonen, M. Haimila and E. Arjas. 2012. Bayesian spatiotemporal analysis of radiocarbon dates from eastern fennoscandia, Radiocarbon 54: 649–659.

Openshaw, S. 1984. The Modifiable Areal Unit Problem, Geobooks, Norwich.

Parnell, A.C., J. Haslett, J.R.M. Allen, C.E. Buck and B. Huntley. 2008. A flexible approach to assessing synchroneity of past *events* using Bayesian reconstructions of sedimentation history, Quat. Sci. Rev. 27: 1872–1885.

Palmisano, A. 2013. Zooming Patterns Among the Scales: a Statistics Technique to Detect Spatial Patterns Among Settlements. pp. 348–356. *In*: G. Earl, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska and D. Wheatley (eds.). CAA 2012: Proceedings of the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Palla Publications, Southampton.

Premo, L. 2004. Local spatial autocorrelation statistics quantify multi-scale patterns in distributional data: An example from the Maya Lowlands, J. Archaeol. Sci. 31: 855–866.

Roberts, Jr., J.M., B.J. Mills, J.J. Clark, W.R. Haas Jr., D.L. Huntley and M.A. Trowbridge. 2012. A method for chronological apportioning of ceramic assemblages. J. Archaeol. Sci. 39: 1513–1520.

Rossignol, J. and L. Wandsnider (eds.). 1992. Space, time, and archaeological landscapes, Plenum Press, New York.

Shannon, C.E. 1948. A mathematical theory of communication. The Bell System Technical Journal 379–423, 623–656.

Shennan, S., S.S. Downey, A. Timpson, K. Edinborough, S. Colledge, T. Kerig, K. Manning and M.G. Thomas. 2013. Regional population collapse followed initial agriculture booms in mid-Holocene Europe, Nat Commun. DOI:10.1038/ncomms3486.

Snodgrass, R.T. 1992. Temporal Databases. pp. 22–64. *In*: A.H. Frank, I. Campari and U. Formentini (eds.). Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Springer-Verlag, Berlin.

Stern, N. 1994. The implications of time-averaging for reconstructing the land-use patterns of early tool-using hominids, J. Hum. Evol. 27: 89–105.

Stoner, W. 2012. Modeling and Testing Polity Boundaries in the Classic Tuxtla Mountains, Southern Veracruz, Mexico. J. Anthropol. Archaeol. 31: 381–402.

Timpson, A., S. Colledge, E. Crema, K. Edinborough, T. Kerig, K. Manning, M.G. Thomas and S. Shennan. 2014. Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. J. Archaeol. Sci. 52: 549–557.

Toizumi, T. and M. Nishino. 1999. Jomonkouki no miyakogawa muratagawaryouiki-kaizukagun. Chibakenbunkazaisentaa KenkyuKiyo 19: 151–171 (In Japanese).

Ullah, I.I. 2011. A GIS method for assessing the zone of human-environmental impact around archaeological sites: a test case from the Late Neolithic of Wadi Ziqlâb, Jordan J. Archaeol. Sci. 38: 623–632.

Yubero-Gómez, M., X. Rubio-Campillo and J. Lopez-Cachero (In press). The study of spatiotemporal patterns in integrating temporal uncertainty in late prehistoric settlements in northeastern Spain, Anthropological and Archaeological Science.

Ziedler, J.A., C.E. Buck and C.D. Litton. 1998. Integration of Archaeological Phase Information and Radiocarbon Results from the Jama River Valley, Ecuador: A Bayesian Approach, Lat. Am. Antiq. 9: 160–179.

Zipf, G.K. 1949. Human Behavior and the Principle of Least Effort. Cambridge: Harvard University Press.