Aoristic Approaches and Voxel Models for Spatial Analysis

Enrico R. CREMA

Institute of Archaeology, University College London, UK.
e.crema@ucl.ac.uk

Abstract

Uncertainties associated with the spatial description of archaeological phenomena are a function of the scale of the analysis. With larger windows of observation we have to deal with an increased degree of ambiguity and imprecision on defining the spatial and temporal dimensions of the phenomena we aim to study. For landscape-scale analysis, these scalar biases are often caused by the subdivision of our evidence into chronological phases with different time lengths, and by the uneven degree of knowledge that we can hope to retrieve from diagnostic artefacts. Temporal uncertainty thus leads to spatio-temporal uncertainty where the pattern we observe at a specific timeframe is a consequence of the definition of the timeframe itself. Traditional GIS approaches based on cartographic representation provide only very poor visualisations of such complex and intrinsically uncertain phenomena. This paper proposes a probabilistic approach for analysing and visualising spatial patterns that addresses, in a quantitative manner, the integration of the temporal uncertainty associated with the archaeological record.

1. Introduction: mapping time

Chiba Southeast New Town is a residential area ca.600 ha in extent in the Boso peninsula, central Japan (*Fig. 1*). The area was investigated during the 1980s and following the tradition of Japanese CRM, large open area excavations were conducted uncovering ca 14% of the entire area. This coverage shed a methodologically consistent light on most of the prehistoric settlements of the Jomon period (ca 13,000–2300 BP), composed by numerous features such as pithouses, storage pits and surface dwellings. Such extensive data allows us to consider individual pithouses as the primary analytical unit and to use the excavation area as a series of explicitly defined sampling windows. From a temporal perspective, Jomon pithouses are usually dated on the basis of a precise relative chronology based on the evaluation of diagnostic pottery. Recently a number of scholars have tried to assign absolute chronological values to these (see Kobayashi 2004) with a high degree of proposed resolution, often reaching a sub-century scale. The dating of features and structures are almost entirely based on such a framework and are very rarely integrated with either absolute dating methods or the temporal topology implied by stratigraphic relationships.

In archaeological terms, the Jomon data from this region is an extremely high resolution dataset, but a problem arises as soon as we try to make use of the temporal dimension by 'mapping' the pithouses associated with each pottery phase. Firstly, the definition of the temporal domain reflects varying degrees of knowledge, based on the quantity and quality of retrieved diagnostic artefacts. Such variation will dramatically reduce our representational capabilities since we have to strike a balance between the desired chronological resolution of our representation and the proportion of our sample that can actually sustain such precision (*Fig. 2*). This reduces our capabilities to represent and analyse any archaeological pattern where the temporal domain is specified, since we are not able to include imprecise data as easily as we might wish. Secondly, the temporal length of each pottery phase is different, thus the direct comparison of the spatial distributions are not feasible, since they are palimpsests formed at different temporal scales. Thirdly, if we wish to enlarge the spatial window of analysis, we have to consider the spatial limits and variability (across time) of chronologies that were defined by pottery styles . And finally, even within the same geographical context, different scholars might propose different temporal subdivisions for the same sequence, producing different spatio-temporal patterns. Therefore, despite an ideal spatial dataset, limitations related to the temporal domain do not
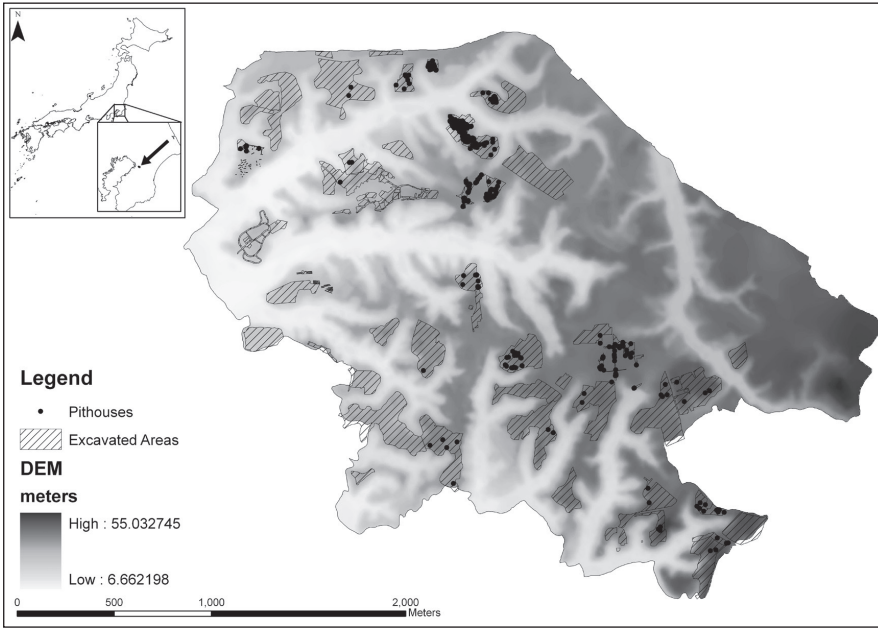

Fig. 1. Location of the Study Area, with the excavation sampling windows and Middle-Late Jomon pithouse distribution.

even allow us to *visualise* properly the available data.

The case presented here is not unique in archaeology and addresses a broader issue of how archaeological spatial patterns are defined. Traditional procedures of mapping and comparing a sequence of temporal snapshots created within a relative chronological framework are clearly not sufficient for any analytical process that might aim to formally integrate the temporal dimension. Different approaches capable of modelling the intrinsic uncertainty and inhomogeneity of knowledge are therefore worth developing.
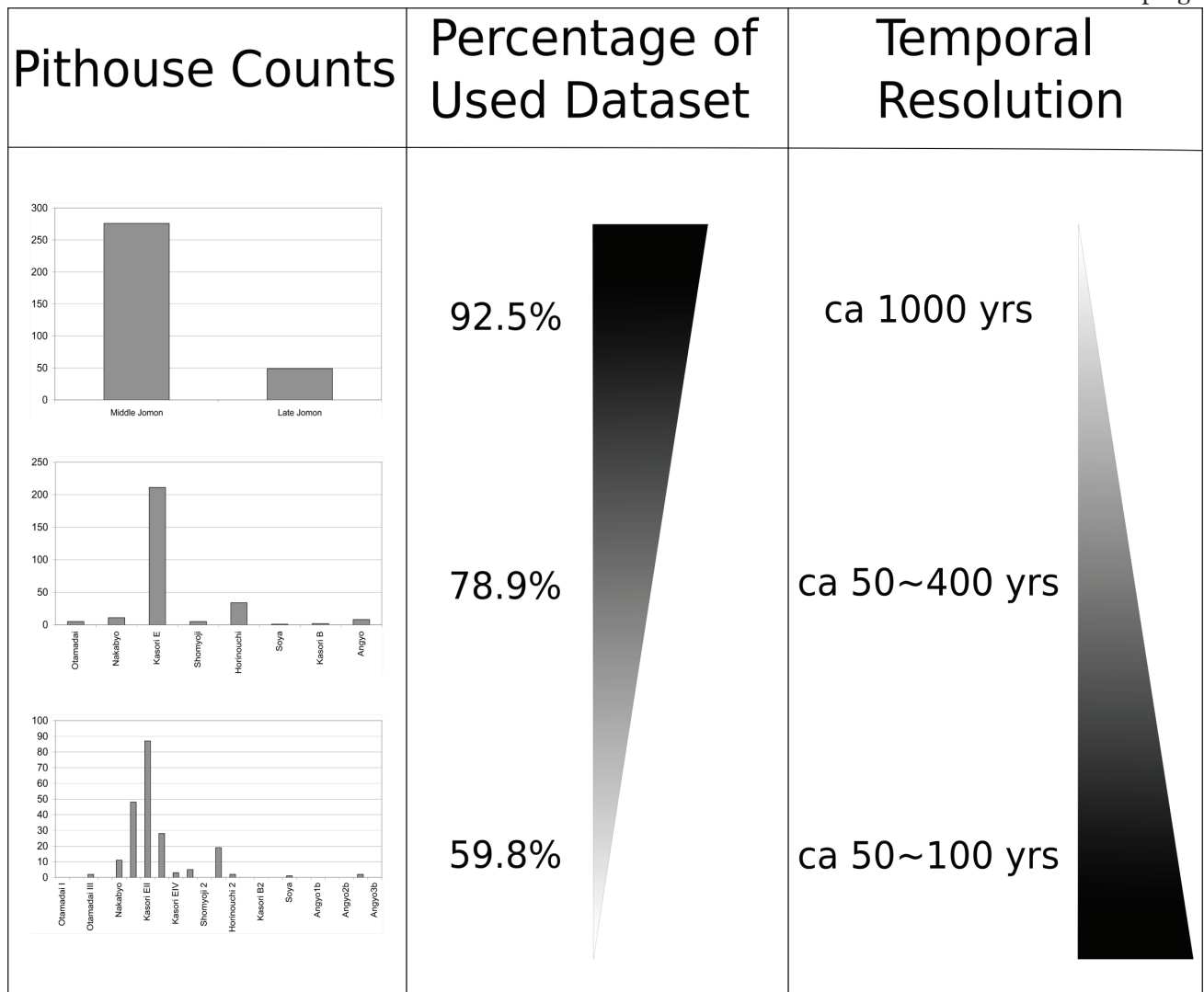

Fig. 2. Pithouse counts time series using different temporal resolutions. Coarser temporal granularity allows the integration of larger portion of dataset, while higher resolution restricts the number of pithouse to be used for the time series.

2. Quantifying uncertainty

The uncertain nature of the archaeological record must be embedded as a quantifiable measure for the purposes of spatial analysis and associated with each spatial unit of analysis, defined as *event* in the statistical jargon. Since we must convey uncertainty when we try to answer the question of '*whether a specific event has occurred or not in a defined moment in time*', probabilistic weighting is the most appropriate way forward. This has been noticed by a number of scholars (see Bevan *et al.* 2008; Johnson 2004; Lock and Harris 2002) who have adopted or proposed different approaches for quantifying the probability of existence of archaeological events.

The quantification of the temporal dimension is considerably affected by a definition of a discrete subdivision in consecutive and sequential series of units (what Snodgrass 1992 called a '*chronon*') that will indicate the *temporal resolution* of the phenomenon we observe. The probability of existence of an event will be negatively correlated with the degree of granularity with which we measure time: lower temporal resolution will produce higher probabilities, and higher resolution will give lower probabilities for each *chronon.*

The probability of existence can be assessed either through the direct retrieval of data related to absolute dating methods or through some form of quantification based on the presence of temporally significant diagnostic artefacts. Both type of information can be successively 'calibrated' through directly non-quantifiable knowledge such as temporal topology or combined by the means of methods based on Bayesian Statistics or Dempster-Shafer Theory.

The abundance of diagnostic artefacts, the heavy reliance of Japanese archaeology towards a temporal definition based on subdivision of pottery phases, and the availability of absolute durations of these has indicated the *aoristic analysis* (Ratcliffe 2000) as the ideal candidate for the probabilistic quantification of the events. The analysis, introduced in archaeology by Ian Johnson (2004) are based on a series of assumptions that must be held in order to be feasible:

1. Each *event* must be definable as a single and unique unit in time and space (e.g. a single feature) rather than a composite and ontologically complex unit (e.g. settlement, site etc.).

2. The duration of the *event* must be instantaneous or at least much smaller than the defined temporal resolution.
3. The duration of all the *events* must be equal.
4. The *terminus ante quem* (*t.a.q.*) and the *terminus post quem* (*t.p.q.*) of each event must be known.
5. The *temporal span* (defined as the temporal length bounded by the *t.a.q.* and *t.p.q.* of an *event* rounded to the precision of the temporal resolution) of the *events* must not be constant. In other words an inhomogenous distribution of temporal knowledge is required.
6. The probability distribution within the temporal span is uniform.

The case study has all these requisites since:
1. the pithouses have been used as a discrete unit of analysis;
2. the duration of a pithouse is considered as relatively short, (maximum of ca. 15 years; Watanabe 1986, 234) in comparison to the temporal resolution adopted (50 years);
3. the duration of the use of pithouses are assumed to be roughly equal; in cases where the re-use of the same pithouse has been recognized, the total aoristic value of the event was greater than 1 and equal to the number of frequentations;
4. the rough duration of the pottery phases in absolute terms is known (Crema 2007) and all the pithouses can be incorporated within this chronological framework;
5. the definition of the relative chronology is based on diagnostic artefacts that provide different degrees of precision ranging from the subphase level ($\approx$50 yrs) to the entire period analysed ($\approx$2000 yrs);
6. the absence of additional sources of temporal knowledge other than the retrieval of diagnostic pottery means that an equal probability for all of the chronons within the time-span is the most straightforward assumption.

Aoristic analysis thus provides a simple and efficient means of assessing the probability of existence for each *event* by attributing an *aoristic weight,* defined as the ratio between the temporal

resolution and time-span of the *event* (see Ratcliffe 2000; Johnson 2004).

Alternative approaches to the probabilistic quantification of temporal uncertainty have also been proposed by Lock and Harris (2002), who, however, focused their work mainly on the development of methods to manage and bind multiple sources of temporal knowledge, without explicitly tackling the issue on how the uncertainty of each event can be quantified.

Aoristic analysis can be interpreted as a special case of probability weighting where the temporal blocks are equally-sized and events do not have duration in time. Even when the first condition is not met, a probabilistic weighting can still be computed for each period as the ratio between its duration and the sum of the durations of all the periods where the event might have occurred. If, with regard to the second condition, events do have duration in time, the probability for each equally long *chronon* can be computed as the ratio between the number of permutations where the events exists in the specific time-block and the total number of possible permutations given a specific time-span and a known duration. The sum of probabilities in this case will be equal to the duration of the event, expressed as a

number of time-blocks. Probability weights can also be computed when both of these conditions are not met, but their computation is more complex and will be treated elsewhere. Finally, in those cases where the type of temporal knowledge involved allows estimation of the shape of the probability function within a time-span (e.g. with calibrated radiocarbon dates), alternative approaches for probabilistic weighting should be considered instead (see Green, this volume).

3. Integrating uncertainty into analysis

The problem of the balance between the chronological resolution and the size of the sub-sample described in *Fig. 2* can be solved by the computation of the sum of aoristic values within each *chronon* (*Fig. 3*). This will provide a finer-grained time series of the population dynamics in probabilistic terms, enabling the detection of previously invisible patterns. The advantage of this method is thus the integration of all the available information, in a discrete and measurable temporal sequence.

The potential analytical framework provided by the aoristic database can be extended for a wide range of statistical methods capable of managing weighted
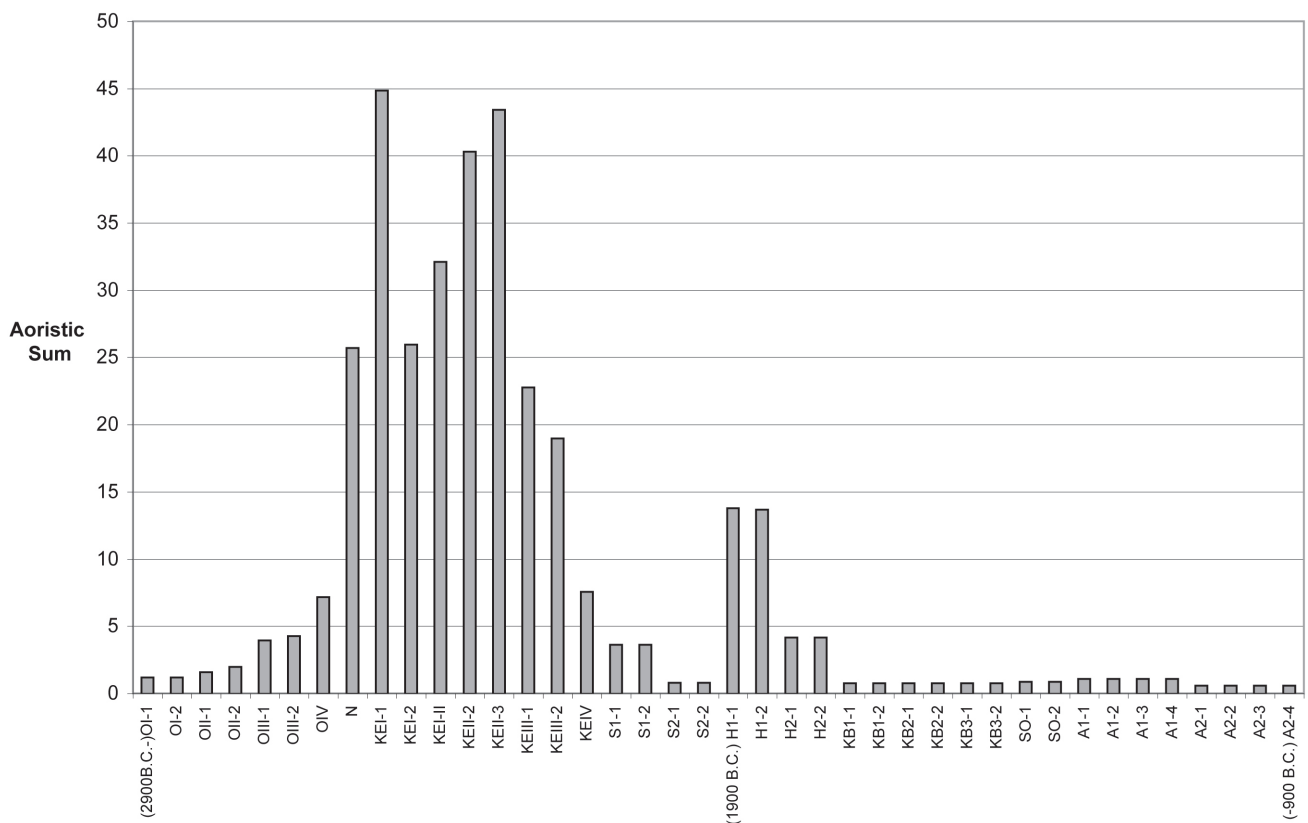
Fig. 3. Aoristic sum of pithouses using a temporal resolution of 50 years. The time-series integrates all the available information using a probabilistic weighting, showing patterns which were previously invisible (cf. Fig. 2).

data. Within the field of spatial analysis, for instance, both the computation of standard distance or the mean centre of distribution of a point dataset have weighted versions where the 'pull' is proportional to the attributed weight. Weighted versions are also available for more advanced techniques such as kernel density estimates (KDE) or Ripley's K functions.

Major advantages of aoristically weighted spatial analysis are the capability of integrating all the available information for each chronon (thus providing a more complete data structure), and the possibility of observing and analysing dynamic changes across time through the comparison of equally long temporal blocks. The latter improvement offers the possibility of adopting a wide range of analytical tools designed for so-called *time series analyses* (TSA). For example, bivariate TSA can assess locations where major change has occurred for each temporal transition, or detect and compare local dynamics in order to assess the synchronicity/ asynchronicity of spatial processes across time.

However, the adoption of these techniques must be treated carefully, since most of the above approaches actually integrate the aoristic weights as *intensities* rather than *probabilities*. While in the assessment of the pattern within each temporal block, this is an acceptable equivalence, problems might rise within a diachronic perspective, where the development of suitable tools are needed to distinguish the lack of knowledge from the lack of pattern. One possible solution, proposed by Crema *et al.* (*in press*), tackles this issue by coupling the probability distribution with temporal Monte Carlo simulation. The core concept of such approach involves the creation of *n* possible spatio-temporal patterns via Monte Carlo simulation, using the probabilistic distribution as the domain of possible permutations. Then each pattern is assessed through standard methods and the results expressed in probabilistic terms, providing a framework for detecting consistent patterns across different spatio-temporal scenarios.

4. Visualising uncertainty

The representation of time within GIS has always been considered as a problematic issue due to the inheritance of the representational framework of traditional cartography. Langran (1992) has shown how time, attribute and location cannot be measured in the same representative framework, and thus only one variable can be quantitatively treated by means of controlling and fixing the other two. Thus for instance, aoristic analysis might be represented as a histogram of aoristic sums (fixed location, controlled attribute, and measured time) or as a sequence of temporal snapshot maps (measured location, controlled attribute, fixed time), etc.

The adoption of a 3-dimensional representation (two spatial and one temporal dimensions) has been conceived as an alternative solution to the problem by a number of authors in different fields, also suggested in archaeology by Lock and Harris (2002). Assuming that the substitution of the third (vertical) dimension is an acceptable reduction of a spatial process, then such a technique allows the representation of the entire spatio-temporal process within a single model, allowing us to locate patterns visually and analytically.

From a practical perspective, the application of real multidimensional representation is still experimental in the social sciences, but an increasing number of specialists in medical and geological fields have started to adopt three dimensional raster models (voxels) for visual and analytical purposes. Their adoption in archaeology is still in its early stages, but the number of applications is increasing, especially in intra-site contexts through the reconstruction of stratigraphic layers (e.g. Bezzi *et al.* 2006).

The actual process by which we construct these cube-like raster models is clearly a critical first issue. Within the 2D domain, the creation of raster models involves either a simple transformation (*rasterisation*) of the data into grid format or the use of an interpolation algorithm. The same principles relevant to rasterisation and interpolation in 2D are also applicable to the 3D case. Voxelisation will thus consist of a *generalisation* of the available information where a series of snapshots are transformed as temporal slices with a vertical Voxel size corresponding to the temporal length of the chronon (Lin 1997). On the other hand, Spatio-temportal interpolation can be implemented using one of a wide number of multidimensional methods such as cubic spline (Mitasova *et al.* 1995) or spatio-temporal kriging (Jost *et al.* 2005). Some interpolators could potentially manage also temporal uncertainty, and Halls *et al.* (2000) for instance proposes a spline interpolator capable of integrating this uncertainty by allowing each event to exert varying 'pulls'. However, the wide range of possible parameter settings required by these interpolators might be difficult to handle and can create multiple alternative

models. Furthermore, most interpolators assume spatio-temporal stationarity and autocorrelation, preconditions which in many cases these are not met in archaeological contexts. The risk in this case is to create models where the space-time between two points are filled by values representing linear change, where abrupt non-linear dynamics might have been present instead .
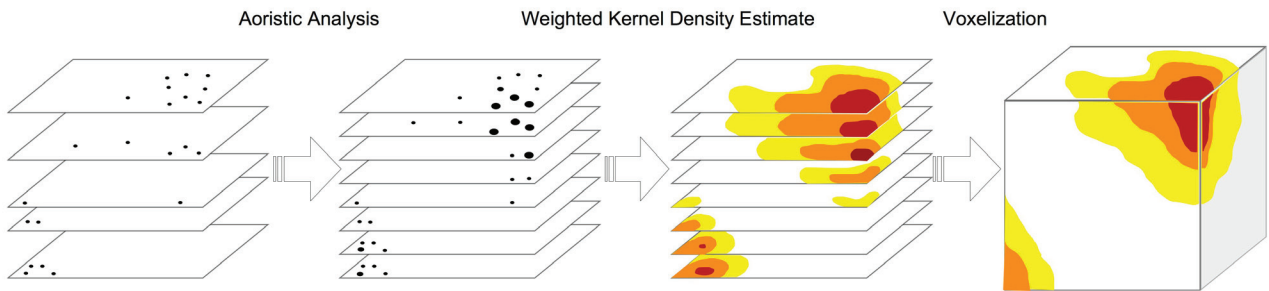

Fig. 4. Workflow for the creation of a Voxel representation of the spatio-temporal process. The data is aoristically weighted, then rasterised through AWKDE and finally converted into a Voxel model.
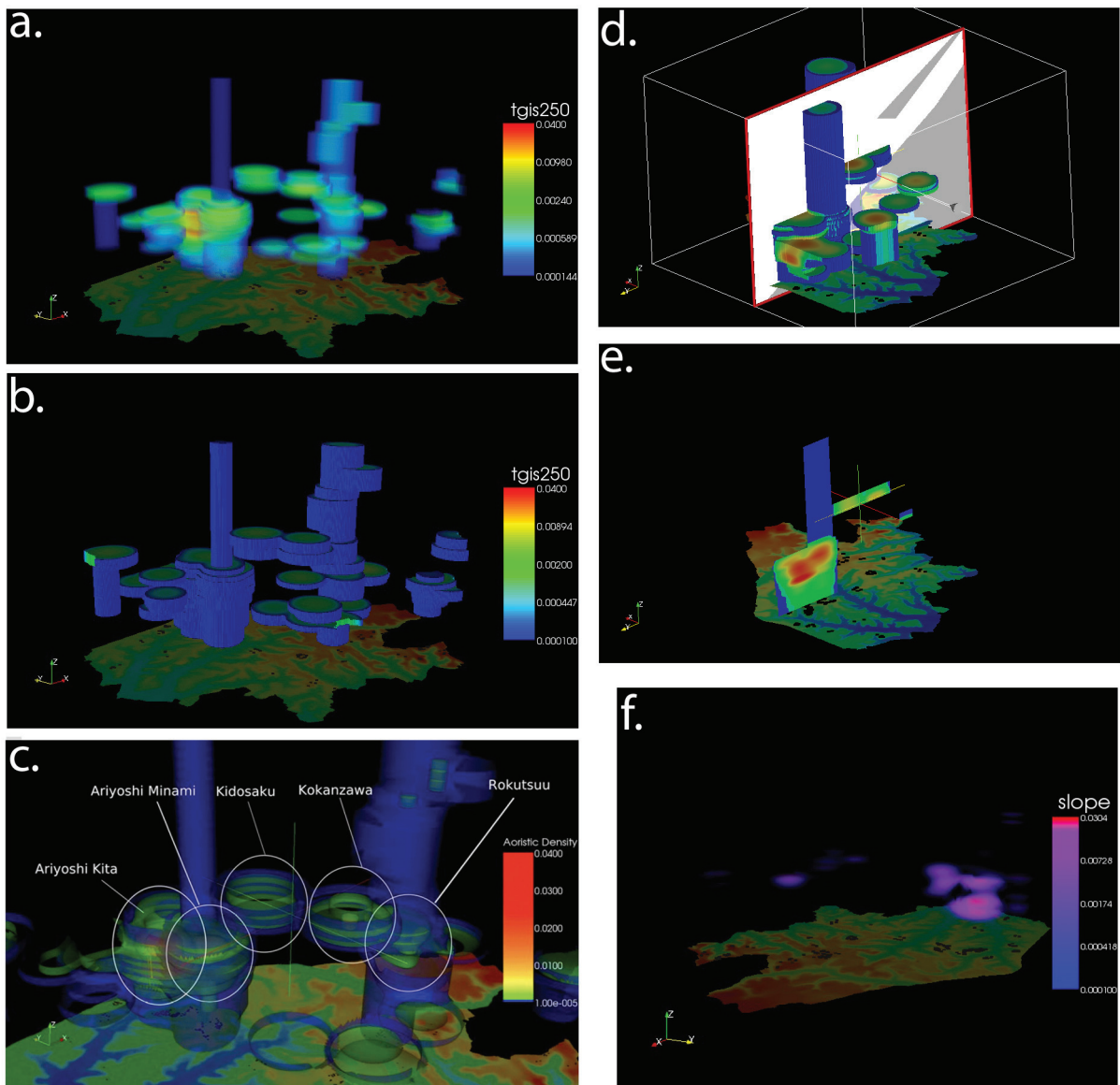

Fig. 5. Voxel representations of the case study area during 2900-900 BC: (a) Volume representation; (b) surface representation; (c) iso-surface representation; (d-e) virtual spatio-temporal cross-section; (f) local density gradient.

The *voxelisation* approach maintains the basic structure of the spatio-temporal information embedding in each cell a numerical attribute value in a specific position defined by a spatial (*x,y*) and temporal (*t*) coordinate system. The adoption of such method for the representation of the probabilistic spatio-temporal distribution is relatively straightforward and might be based either on the direct rasterisation and the successive voxelisation of the pithouses or to the voxelisation of some raster surface, such as the aoristically weighted KDEs (*Fig. 4*). For the present study, the latter approach has been adopted, using the *r.to.rast3* module in GRASS GIS 6.3 (http://grass.itc.it/) for the ordered sequence of the density maps. The voxel has been created using an x and y resolution equal to the original 2D raster maps of the area (5 meters) and with a vertical resolution of 50 yrs. The definition of the upper and lower limit of the voxel model (representing the temporal boundaries) has been previously set using the absolute chronologies in negative values (bottom=-2900; top=-900) with the *g.region* command. Since GRASS does not allow a direct representation of voxel models, the entire dataset has been converted in .vtk format (using the *r3.out.vtk* command) and imported into ParaView, an open-source software package for parallel visualisation that is capable of handling voxel data (http://www.paraview.org/).

ParaView provides three different ways to visualise the voxel model, namely *volume*, *surface* and *iso-surface*. The first provides a blurry representation of the entire dataset in a series of semi-transparent cells (*Fig. 5a*), the second is based on a solid representation (thus with only the 'external' values visible), (*Fig. 5b*) and the third might be considered as a three-dimensional version of contour maps, with a user- defined interval (*Fig. 5c*). ParaView also allows the creation of virtual cross sections in time space (*Fig. 5d–e*) or the extraction of defined intervals of values to create series of subset models. Possible extensions of spatio-temporal map algebra have also been proposed by various authors (Lin 1997; Mennis *et al.* 2005); however, despite different approaches of focal and zonal calculations, the definition of an optimal multidimensional neighbourhood is still an open debate. Even so, multidimensional map algebra allows us to compute derivative analytical surfaces, based on local computations, such as low and high pass filters, or three dimensional gradient maps (e.g. *Fig. 5f*).

5. Conclusion

Aoristic analysis and voxel models can thus be used to visualise a spatio-temporal process in a single representational framework where the intrinsic uncertainty of archaeological data is retained. However, voxel representations are still in an experimental phase and clear limitations are still evident. The model is capable of handling and representing single attribute at a time, and considering that the probabilistic values are mandatory for the creation of temporal slices, extending the representational framework to other variables will require the creation of separate models where the probabilistic values are embedded separately. From a computational perspective, this is clearly not problematic, and the availability of zonal map algebra could already provide basic query and analytical tools across different models however from a visual perspective, the adoption of voxel overlays will decrease dramatically the ease with which we would then be able to perceive and understand the spatio-temporal processes. Furthermore, the basic on-the-fly query functionality available for most GIS packages are still not present in 4D, and thus distances in space and time cannot be measured directly, and spatial and temporal scale cannot be shown on the three-dimensional map. From analytical perspective, multidimensional map algebra is still rare in geography, and it is therefore, in contrast, the current availability of a larger number of time series analysis tools that has more immediate possibilities for archaeology. Thus, at this stage, the advantage of three-dimensional voxel model is primarily that it provides a general overview of a spatio-temporal process, while the main analysis is still conducted within a traditional cartographic framework of sequential snapshots, albeit underpinned by statistical formality. The development of proper spatio-temporal analysis through the quantification of uncertainty however is a good starting point from which to develop a more formal and complete temporal GIS.

6. Acknowledgments

References

Bevan, Andrew, James Conolly, and Aris Tsaravopoulos (2008). The Fragile Communities of Antikythera. *Archaeology International* 11.

Bezzi, Alessandro, Luca Bezzi, Denis Francisci and Rupert Gietl (2006). L'utilizzo di voxel in campo archeologico, *Geomatic workbooks* 6. (http://www.arc-team.com/Upload/Geom_Workbooks6_voxel_archeologia_it.pdf).

Crema, Enrico R. (2007). *Temporal Uncertainty in Spatial Analysis: a case study on meso-scale settlement pattern analysis of Middle and Late Jomon period in Chiba, Japan.* Unpublished Master's Dissertation. Deposited at the Institute of Archaeology, University College London, London.

Crema, Enrico R., Andrew Bevan and Mark Lake (*in press*). A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record. *Journal of Archaeological Science.*

Green, Chris (*in press*). It's about time: temporality and intra-site GIS. *Proceedings of the Computer Applications in Archaeology Conference, Budapest 2–6 April 2008.*

Halls, Peter J., Fiona A. C. Polack and Simon E. M. O'Keefe (2000). A new approach to the spatial analysis of temporal change using todes and neural nets. *Cybergeo* 139 (http://www.cybergeo.eu/index911.html).

Johnson, Ian (2003). Aoristic Analysis: seeds of a new approach to mapping archaeological distributions through time. In: Karin F. Ausserer, Wolfgang Börner, Maximilian Goriany, and Lisa Karlhuber-Vöckl (eds.) *[Enter the Past] The E-way into the Four Dimensions of Cultural Heritage: CAA2003*, Vienna: BAR International Series [CD-ROM version]. 448–452.

Jost, G., G. B. M. Heuvelink and A. Papritz (2005). Analysing the space–time distribution of soil water storage of a forest ecosystem using spatiotemporal kriging. *Geoderma* 128, 258–273.

Kobayashi, Ken'ichi. (2004). *Jomonshakaikenkyu no shishiten: tanso14nendaisokutei no riyo* [New perspective of study on Jomon society: Applied radiocarbon dating to Archaeological Methodology], Tokyo:Rokuichishobo.

Langran, Gail. (1992). *Time in Geographic Information Systems.* London: Taylor and Francis.

Lin, H. (1997). STIN operation: A study on searching for spatiotemporal measures. *Cartography* 26, 27–36.

Lock, Gary and Trevor Harris (2002). Analysing change through time within a cultural landscape: conceptual and functional limitations of a GIS approach. In: Paul Sinclair (eds.) *The Development of Urbanism from a Global Perspective.* Uppsala: Uppsala Universitet. Retrieved 10th November 2007 from World Wide Web: (http://www.arkeologi.uu.se/afr/projects/BOOK/contents.htm).

Mennis, Jeremy, Roland Viger and Dana C. Tomlin (2005). Cubic Map Algebra Functions for Spatio-Temporal Analysis. *Cartography and Geographic Information Science* 32, 17–32.

Mitasova, Helena, Lubos Mitas, William M. Brown, David P. Gerdes, Irina Kosinovsky and Terry Baker (1995). Modelling spatially and temporally distrubuted phenomena: new methods and tools for GRASS GIS. *International Journal of Geographical Information Systems* 9, 433–446.

Ratcliffe, Jerry H. (2000). Aoristic analysis: the spatial interpretation of unspecific temporal events. *International Journal of Geographic Information Science* 14, 669–679.

Snodgrass, Richard T. (1992). Temporal Databases. In: Andrew H. Frank, Irene Campari and Ubaldo Formentini (eds.) *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space.* Berlin: Springer-Verlag, 22–64.

Watanabe, Hitoshi (1986). Community Habitation and Food Gathering in Prehistoric Japan: An Ethnographic Interpretation of the Archaeological Evidence. In: Richard J. Pearson, Gina L. Barnes and Karl L. Hutterer (eds.) *Windows on the Japanese Past: Studies in Archaeology and Prehistory.* Ann Arbor: Centre for Japanese Studies University of Michigan, 229–254.