BOREAS An international journal of Quaternary research

Spatial simulation and modelling of the early Pleistocene site of DS (Bed I, Olduvai Gorge, Tanzania): a powerful tool for predicting potential archaeological information from unexcavated areas

MANUEL DOMÍNGUEZ-RODRIGO, LUCÍA COBO-SÁNCHEZ, DAVID URIBELARREA, MARÍA CARMEN ARRIAZA, JOSÉ YRAVEDRA, AGNESS GIDNA, ELIA ORGANISTA, AINARA SISTIAGA, DAVID MARTÍN-PEREA, ENRIQUE BAQUEDANO, JULIA ARAMENDI AND AUDAX MABULLA

BOREAS

Domínguez-Rodrigo, M., Cobo-Sánchez, L., Uribelarrea, D., Arriaza, M. C., Yravedra, J., Gidna, A., Organista, E., Sistiaga, A., Martín-Perea, D., Baquedano, E., Aramendi, J. & Mabulla, A.: Spatial simulation and modelling of the early Pleistocene site of DS (Bed I, Olduvai Gorge, Tanzania): a powerful tool for predicting potential archaeological information from unexcavated areas. *Boreas*. 10.1111/bor.12252. ISSN 0300-9483.

Spatial statistical models are powerful tools for creating simulation and prediction models. Here, we apply such models to the newly discovered 1.84 Ma site of DS (Bed I, Olduvai Gorge, Tanzania). Ongoing excavation has already exposed 370 m^2 of the same discrete archaeological level. This is the biggest window into an Early Pleistocene anthropogenic site. With such a large area opened, modelling based on spatial trends (using coordinates) and on covariates (topography) has enabled the creation of predictions about where the densest concentrations of unexcavated materials may lie. Following this modelling, excavation has confirmed the predictions; the densest clusters of stone tools and fossils bones are palaeotopographically and palaeoecologically influenced. Spatial statistical analysis is, therefore, a powerful analytical tool to model and understand in-site and off-site hominin behaviour as an interaction between hominins and environments.

Manuel Domínguez-Rodrigo (m.dominguez.rodrigo@gmail.com), Lucía Cobo-Sánchez, David Uribelarrea, María Carmen Arriaza, José Yravedra, Elia Organista, Enrique Baquedano and Julia Aramendi, Institute of Evolution in Africa (IDEA), Covarrubias 36, Madrid 28010, Spain; Manuel Domínguez-Rodrigo, Real Colegio Complutense at Harvard, 26 Trowbridge Street, Cambridge, MA 02138, USA; Manuel Domínguez-Rodrigo, Lucía Cobo-Sánchez, José Yravedra and Julia Aramendi, Department of Prehistory, Complutense University, Madrid 28040, Spain; David Uribelarrea and David Martín-Perea, Department of Geodynamics, Complutense University, Madrid 28040, Spain; María Carmen Arriaza, School of Geography, Archaeology and Environmental Studies, Evolutionary Studies Institute, University of Witswatersrand, Private Bag 3, Johanesburg 2050, South Africa; Agness Gidna and Audax Mabulla, Paleontology and Archaeology Unit, National Museums of Tanzania, Dar es Salaam, Tanzania; Ainara Sistiaga, Geobiology and Astrobiology Lab, EAPS Massachussetts Institute of Technology, 45 Carleton St. E25-647, Cambridge, MA 02139, USA; Enrique Baquedano, Museo Arqueológico Regional, Plaza de las Bernardas, Alcalá de Henares, Madrid, Spain; *received 20th October 2016, accepted 9th March 2017.*

Clarke, one of the founders of spatial archaeological research, argued that human behaviour was spatially organized (Clarke 1977, 1979). One of the advantages of a spatial approach to archaeology is to uncover patterns that could convey behavioural information. These patterns, resulting from non-random human decisions, are the result of regularities in the spatial materialization of those behaviours. Although spatial archaeology has a long tradition of application in the 20th century, most of its development and use has been graphic rather than statistical. This has prompted a wide theoretical use of spatial archaeology to elaborate on social interpretations, not necessarily accompanied by powerful statistical apparatus (Hodder 1977a; Ashmore 2002, 2014). However, knowledge of statistical approaches to spatial analyses in archaeology has long been available (Hodder & Orton 1976; Hodder 1977b; Donnelly 1978; Brughmans 2012; Wurzer *et al.* 2014). Most commonly, widely used statistical analyses involve simple nearest-neighbour and distance approaches, which differ from the complex multivariate statistics that are currently applied in other archaeological subfields (e.g. Arriaza & Domínguez-Rodrigo 2016).

Frequently, in the graphic display of spatial analyses, the detection of any associations is left to the eye of the observer. However, associations are frequently much more abundant and subtle than subjective naked-eye appreciations can find. When combining spatial analysis within a multivariate statistical framework, inferences are more solid. A body of spatial statistical tests has been developed within the fields of geology, epidemiology, ecology and econometrics (e.g. Pebesma 2004; Baddeley & Turner 2005; Bivand 2010; Roger *et al.* 2013; Dorman 2014; Baddeley *et al.* 2015). These powerful tests are essential tools for spatial archaeology. A sound statistical approach to archaeological spatial analyses can contribute to understanding behaviour through spatial patterns, as argued by Clarke (1977), and to making predictions on archaeological information that has either not been found or not been excavated yet. For example, good statistical regression models and simulations can help detect the most suitable areas for preservation of specific archaeological sites, which can help in planning surveys, or they can help to estimate which unexcavated areas are more likely to contain certain types of archaeological information. They can also

DOI 10.1111/bor.12252

contribute to understanding the effects of multiple variables in the spatial configuration of patterns and how these relate to human behaviour.

Here we will show an example of the utility of statistical spatial analysis for selecting the potentially densest areas of an unexcavated portion of a site. This was achieved by prediction stemming from regression modelling. This approach was applied to a newly discovered early Pleistocene site (DS) at Bed I in Olduvai Gorge (Tanzania). Predictions obtained through modelling were subsequently tested and confirmed. This opens the door to the restoration of statistical spatial analysis as an important analytical tool in archaeology. It also stresses its relevance to the study of postdepositional processes affecting archaeological assemblages as well as behavioural patterns that are ultimately responsible for the configuration of those assemblages.

Material and methods

DS (David's Site), Bed I (Olduvai Gorge)

DS was discovered by The Olduvai Paleoanthropology and Paleoecology Project (TOPPP) in 2014. It was located in an area covered by a dirt road that had been used since M. Leakey's early exploration of the gorge more than 50 years ago. For generations, archaeologists have walked along that road while traveling to classical sites, such as HWK. In 2014, the rains had started eroding the archaeological deposit, which was subsurficial and therefore, strongly exposed to subsequent erosion (Fig. 1). It was necessary to excavate the larger exposed area to retrieve as much information as possible before erosion seriously damaged a greater extension of the assemblage. This changed TOPPP's original plans, and in 2014–16, all efforts were focused on exposing as large an area as possible on the erosion-affected platform. This was also accomplished because a large digging crew (ranging between 34 and 58 team members depending on the year) was available. An area of 370 m^2 was initially opened. This area, bigger than that excavated at FLK Zinj (Leakey 1971) (until now, the biggest open window to the African Early Pleistocene), exhibited a dense concentration of fossil bones and stone tools (Fig. 2). The materials were uncovered through careful excavation, involving sieving and plotting with laser theodolites. Once the palaeosurface containing the discrete archaeological level (<10 cm deep) was uncovered, it was stereo-photographed to obtain a photogrammetic 3D reconstruction. Then, the azimuth and plunge of artefacts and fossil bones were measured. Subsequently, plotting and removal of items from the ground took place. On-going analysis of the materials shows a high integrity of the site. Refits are numerous and several articulated or semi-articulated anatomical portions have been unearthed. The abundant and diverse stone tool kit is functionally associated with the thousands of bones discovered. These are mostly green-broken, and both percussion and cut marks have been documented on a significant portion of the bone assemblage. Impact flakes are abundant, and at the moment of writing, only two long bone ends bearing carnivore furrowing have been discovered. This would initially suggest a minimal impact of carnivore damage, which would explain the abundance of axial remains, namely rib specimens. The site, given its size and highly anthropogenic signature, will certainly be a great addition to our understanding of early hominin behaviour during Bed I times.

Statistical modelling

A Poisson point process is a random spatial distribution of points that are independent from one another (Cressie 2015). When points are stationary (their density is proportional to area) they display a homogeneous pattern. A spatial trend is documented when intensity is conditioned by location, thereby creating inhomogeneity. The intensity of inhomogenous patterns can be understood by a log-linear function of the covariates (Baddeley *et al.* 2015). The principle of maximum entropy, whereby all possible spatial distributions are considered, is the basis for calculating the intensity resulting by constraints in the observed data. Regression techniques are commonly used to detect these influencing spatial dispersal and concentration processes (Bivand *et al.* 2013). Regressions allow the fitting of different models and detect which variables have a statistically significant influence on the intensity of points. Complete spatial randomness (CSR) (also known as uniform Poisson process) depends on point independence and homogeneity. Statistical methods are available for testing CSR, such as chi-square and Kolmogorov–Smirnov tests. If any given point pattern supports CSR, log-linear methods can be used parametrically (e.g. linear regression) to understand homogeneous and inhomogeneous Poisson point processes. If CSR is rejected, then alternative non-parametric methods (e.g. polynomial regressions) are required (Baddeley *et al.* 2015).

Nonstationary Poisson models can be fitted via polynomial regressions, which use a log-quadratic intensity model (Baddeley *et al.* 2015; Oyana & Margai 2015). In some cases, inhomogeneous models can be fitted with intensity estimates that are proportional to covariates. This is referred to as a baseline model and logarithmic transformations are required. This regression type creates models with offsets. When interdependence of points is either suspected or documented, such a positive dependence leading to clustering can be best approached via Cox and Cluster models (Baddeley *et al.* 2015). These models are modifications of the Poisson process by incorporating random effects, which enable the capture of the intricacies of inhomogeneous and interdependent point processes.

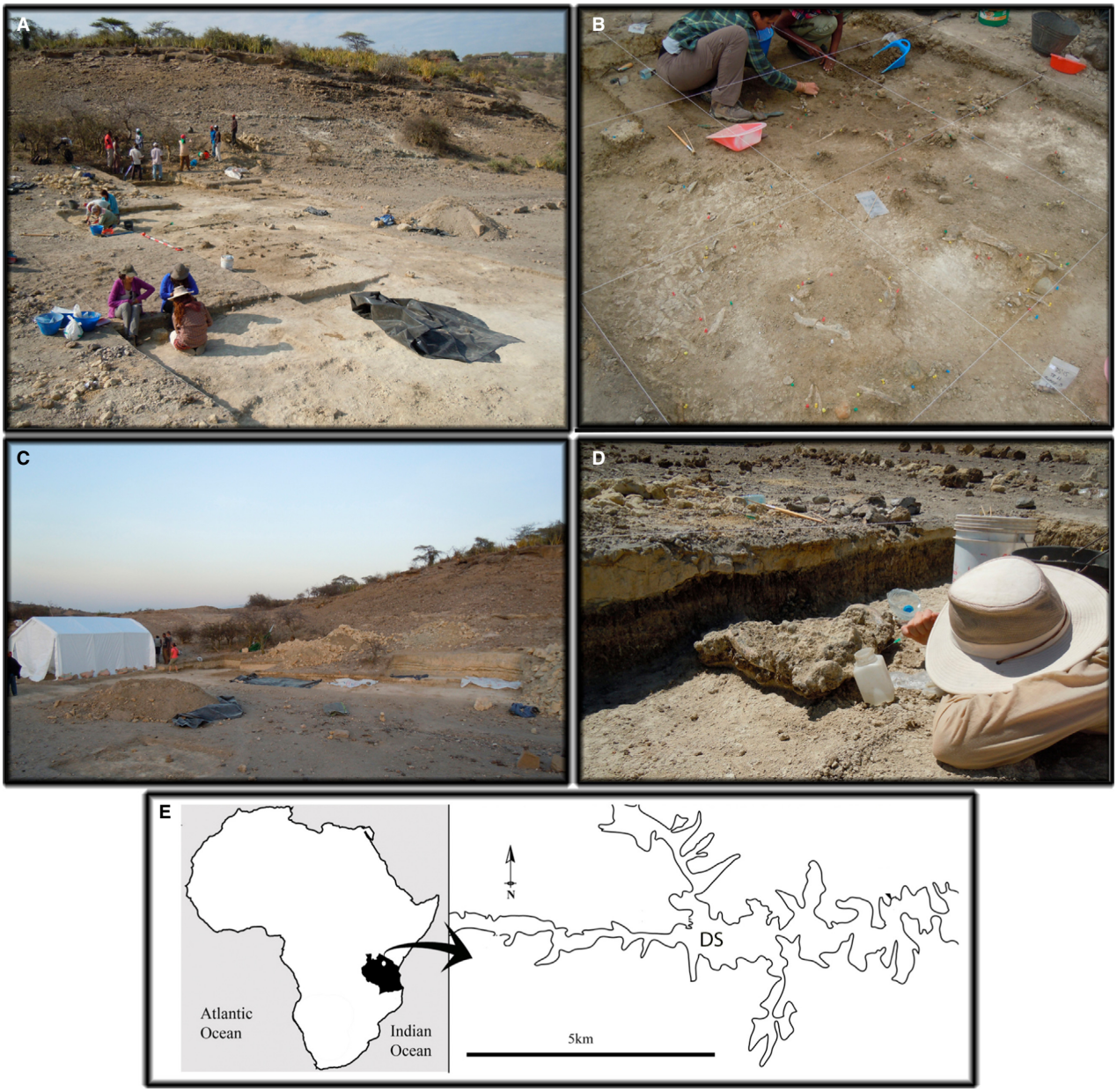

Fig. 1. A. General view of the DS excavation looking south. B. Detail of the spatial distribution of materials in one of the trenches. C. Southern view of the excavation toward the slope, with the forensic unit to the left for micro-residue analysis. D. Detail of the excavation of a kudu skull with the stratigraphical reference of the double-coloured 'Zinj' clay stratum overlaid by the Tuff IC. E. Insert with location of Tanzania and Olduvai Gorge and situation of DS.

In the present study, we plotted the archaeological materials within the excavated area of 370 m^2 (window) against the density map generated by the intensity of points. Correction methods to avoid the window edge effect were applied. Diggle's algorithm was applied to minimize the mean-square error (Diggle 2003). A Cox process was assumed given the clustering pattern documented (see below). This method helped to define the best sigma option (bandwidth) to get an accurate density map. To test the CSR null hypothesis of the point pattern, a chi-square test was used. In order to do this, the excavated area was artificially divided into a 2×2 m grid. The test showed that the pattern was not a homogeneous Poisson point process ($\chi^2 = 4276.783$, df = 109, *p*-value < 2.2e-16). The K and L functions indicate a clustering spatial trend (Fig. 3A, B). The functions were created by combining standard K and L tests with edge correction and envelopes with 95% confidence intervals, created via Monte Carlo simulation (n = 39) (Baddeley *et al.* 2015). Given that the K and L functions assume a homogeneous Poisson pattern, we also used the K and L functions adapted for inhomogeneous patterns and with

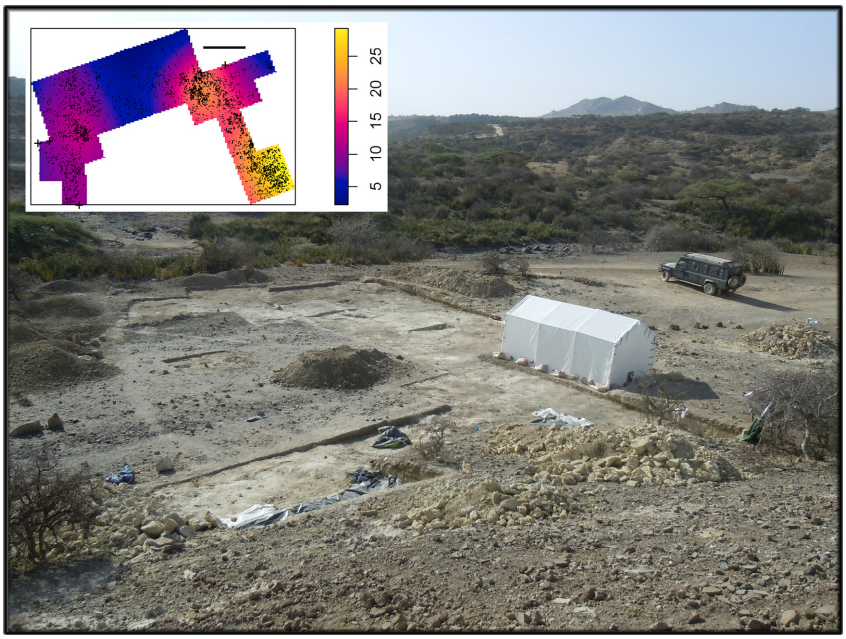

Fig. 2. General view of the DS excavation taken from the south. Insert: map of distribution of materials (bones and stone tools) from the main archaeological layer. Colour indicates different densities per unit. Scale bar = 6 m.

edge correction (including Diggle's sigma selection) (Diggle 2013; Baddeley *et al.* 2015). These tests show a clustering trend composed of clusters of small radii, outside of which the pattern is either random (K) or slightly clustered (L) (Fig. 3). The L function may better illustrate the higher intensity documented to the south of the excavation when compared to the east and west. A Clark–Evans test of aggregation of a point pattern was performed to further support the interpretations from the previously described functions (Clark & Evans 1954). A Monte Carlo approach was used by using 100 replications. The alternative hypothesis is that $R < 1$ corresponds to a clustered point pattern. The result rejected the null CSR hypothesis and supported the alternative hypothesis that the assemblage is moderately clustered ($R = 0.733$, $p = 0.01961$).

Subsequently to these tests, a new window (50×30 m) was created surrounding and including the excavated area, but mostly composed of unexcavated terrain. This rectangular window was used to make predictions about a broader spatial point pattern trend than that documented within the excavated window. This was attempted through regression methods, using the 'ppm' function of the R 'spatstat' library (Baddeley *et al.* 2015). First, a linear regression model was used. Subsequently, given the inhomogeneous nature of the excavated materials at DS, a square polynomial regression was used. In both cases, the resulting models were used to predict directionality of the spatially varying intensity.

The inhomogeneous nature of the DS point pattern suggested that the points were not independent (as required by Poisson process modelling). Bones and stone tools seemed functionally associated, and therefore, their spatial distribution is not independent. Likewise, the abundant refitting found (on-going research) shows that the spatial distribution of element NISPs is also dependent on element types. The inferred degree of dependency suggested that regressions should consider cluster processes, which could include point interdependence. For this reason, Cox–Cluster regression models (via the 'kppm' function of the 'spatstat' R library) were used (Guan 2006; Waagepetersen 2007). These models were the baseline for a series of simulations following the regression results. These simulations were undertaken using two methods: Thomas and Log Gaussian Cox Process (LGCP) (Møller *et al.* 1998). Cluster processes using the Thomas method focus on cluster formation, and this is simulated by a two-step process. In the first step, a pattern of 'parent' points is generated. Subsequently, each 'parent' point generates a random pattern of 'offspring' points. Only the latter is considered (Baddeley *et al.* 2015). In the Thomas cluster process, the offspring clusters following an isotropic Gaussian density pattern. The Thomas process is 'a Cox process, with random driving intensity A(u) equal to the superposition of Gaussian densities centred at each of the parent points' (Baddeley *et al.* 2015: p. 463).

A LGCP is a Cox process based on random fields, where intensity is measured by: $A(u) = \exp G(u)$, where A (u) is the driving intensity and G(u) is a Gaussian random field (Møller *et al.* 1998). These random fields are simulated processes with spatially varying intensity, yielding hot (dense) and cold (scatter) point areas. Given the combination of cluster-scatter point patterns repli-

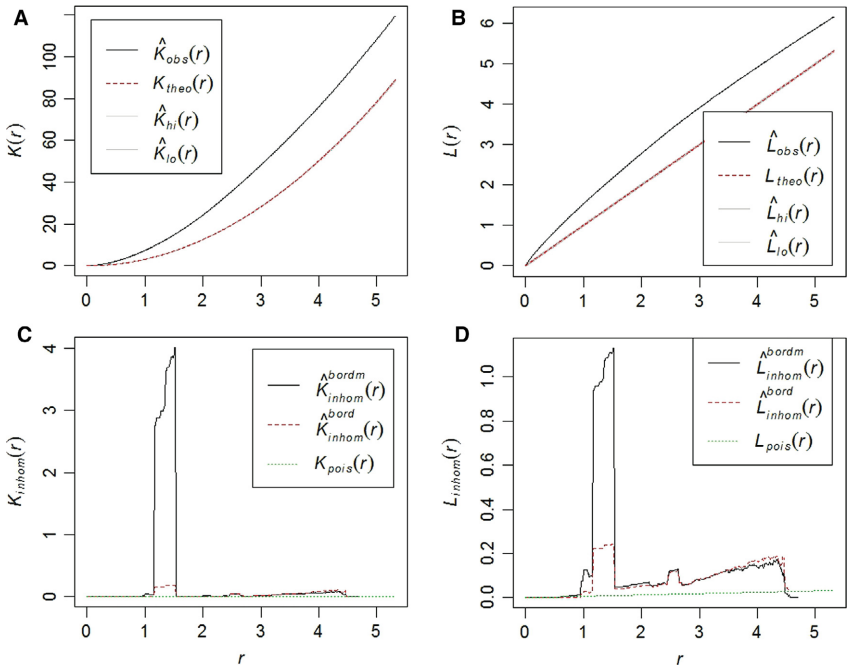

Fig. 3. A and B. K and L functions with edge correction and envelopes with 95% confidence intervals, created via Monte Carlo simulation. C and D. K and L functions adapted for inhomogeneous patterns and with edge correction (including Diggle's sigma selection).

cated and simulated in the random fields, this method more realistically reproduce point pattern models that one may find in a random spatial distribution of archaeological points.

Both the Thomas and LGCP simulation processes operate with intensity ranges and spatial distributions as observed in the sample used for elaborating the model. The higher degree of clustering and scattering (depending on the method) takes the original sample distribution and size into consideration to project the simulated sample in similar terms either inside the original window or in any other window in which the same spatial patterning is reproduced (Baddeley *et al.* 2015). Thus, a projection of predicted spatial intensity within a new window does not intend to faithfully reproduce the net real intensities in that window, but their spatial trend, provided the original window from which the model is obtained is representative of the intensity trend within a broader framework.

Once the spatial trend was captured and modelled using the x-y coordinates as the covariates, we tested if the distribution of archaeological materials was correlated to the palaeosurface topography. If the topography detected elevations and depressions, a preferential distribution of materials on depressed areas could be indicative of postdepositional disturbance. If materials were deposited on higher ground, this could be used to infer a smaller impact of postdepositional re-sedimentation processes, as well as contribute to model predictions on material distribution in the unexcavated window. Covariate correlation was measured via the Kolmogorov–Smirnov test and the Berman test (null hypothesis is CSR) (Berman 1986). Covariate effect was tested with the estimation of parameter p (probability of density) using a resource selection function, built with

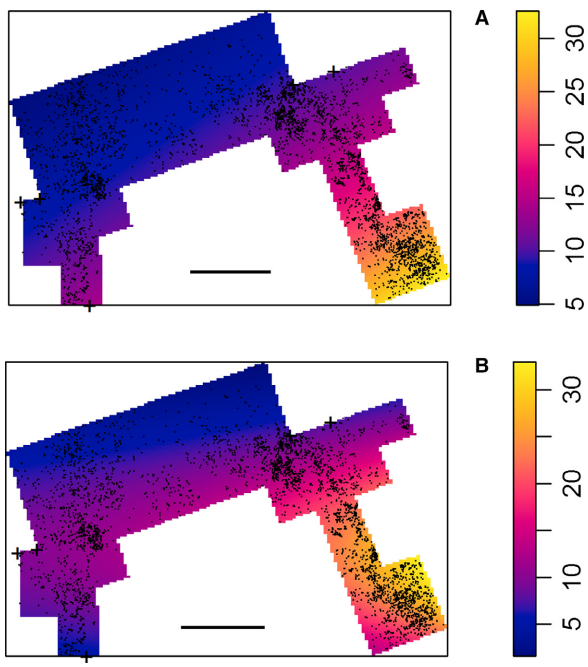

Fig. 4. Distribution of the excavated materials plotted against density estimates (in colour) obtained through linear (A) and polynomial (B) regressions. Scale bars = 6 m.

kernel smoothing. Simple linear regression was subsequently used to predict the distribution of archaeological materials according to the palaeosurface topography. Once fitted, the prediction was tested against the documented distribution of materials. This validation test compares the differences between predicted intensity and observed intensity. This was done using the 'eval.im' function of the 'spatstat' R library, which produces a result of 0 if the prediction matches the documented intensity and a negative number if there is a disagreement in the predicted and documented intensities (Baddeley *et al.* 2015).

A linear regression and a polynomial regression were carried out using the topography of the palaeosurface as a covariate to predict intensities within the excavated area and test simulation models. The simulation models were also created with the LGCP random field algorithm, as above (Møller & Waagepetersen 2003). A control model was constructed using an independent framework, in which the 'ppm' function was used. The results were very similar to the regression using the Cox–Cluster model. This control regression enabled us to validate the model by comparing the model to the data. Important regression diagnostics are leverage and partial residuals. These were analysed by using point process residual measures, smooth residual fields and lurking variable plots on the x-y coordinates (Baddeley *et al.* 2015).

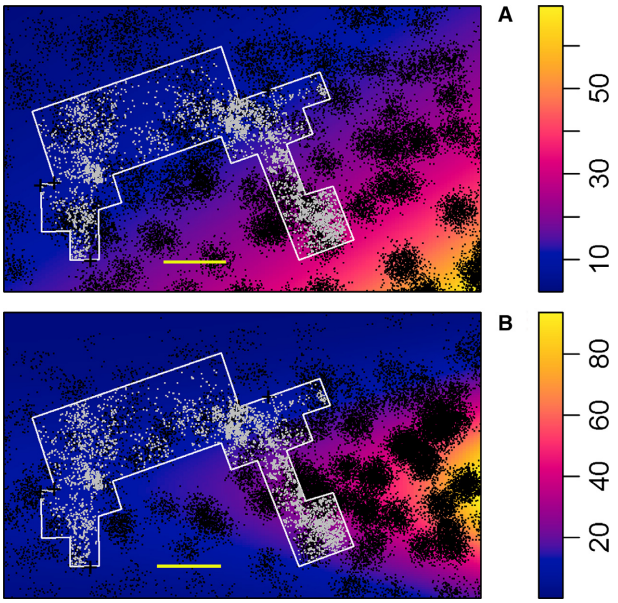

Fig. 6. Distribution of the excavated materials at DS (light grey dots) plotted against estimates of material distribution as predicted by regressions (black dots) and compared to density estimates (in colour and black dots) for the unexcavated surrounding space. This was obtained through linear (A) and polynomial (B) regressions. In both cases a clustering Thomas algorithm was used (Baddeley *et al.* 2015). Scale bars = 6 m.

Archaeological testing of the model

Following the indications of the simulation models (Fig. 7), several areas were selected as testing units for these models. The three main ones occur adjacent to the southern trenches in the excavated area and are referred to as 'a', 'b' and 'c'. These are special testing units, because (i) they will confirm or reject if the models are correct and (ii) they will show which model had a higher capacity of prediction. As a complement, to complete the testing of the models, a trench ('d') was excavated to the east of the excavated area. The hypothesis to be tested (following the indications of the models; see below) is that intensity will decrease following this trench order: b-c-a-d.

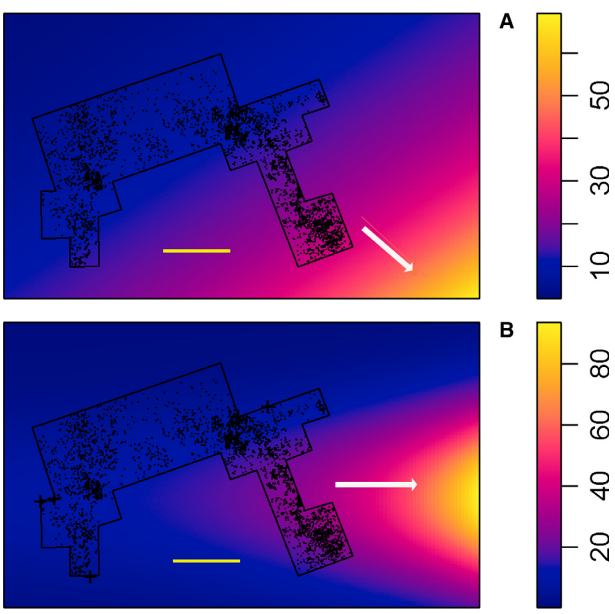

Fig. 5. Distribution of the excavated materials at DS plotted against density estimates (in colour) for the unexcavated surrounding space obtained through linear (A) and polynomial (B) regressions. Arrows indicate trend of increasing density. Scale bars = 6 m.

Excavations were conducted by opening three 6×3 m trenches (a–c) and a smaller 3×3 m trench (d). These units allowed accurate and timely documentation via photogrammetry and laser total stations without exposing the materials for too long to subaerial conditions. Each lithic artefact and fossil bone was recorded in its exact location and orientation. Densities were evaluated per square metre units. A comparison between the predicted and the obtained values was made using correlations of raw estimates.

Results

Selection of models

Regressions (linear and polynomial) fitted to the excavated window coincide in showing that the spatial trend is expressed as an increase in intensity towards the south of

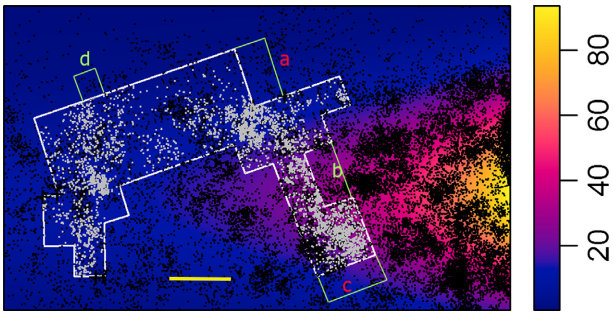

Fig. 7. Distribution of the excavated materials at DS (light grey dots) plotted against estimates of material distribution as predicted by regressions (black dots) and compared to density estimates (in colour and black dots) for the unexcavated surrounding space. This was obtained via a square polynomial regression using the LGCP method with a random field algorithm (Baddeley *et al.* 2015). The testing trenches are indicated (a, b, c, d). Scale bar = 6 m.

the excavation (Fig. 4). The polynomial model fits better the actual distribution of materials, as expected given the inhomogeneous nature of the point pattern. This is especially noticeable towards the east of the excavation. The extension of the model to the delimited area of the big window (comprising the area surrounding the excavated window) shows two different trends of increased intensity depending on the model (Fig. 5). The linear model suggests an increase of intensity towards the west. The polynomial model suggests an increase of intensity towards the south.

A set of Cox regressions (using both linear and polynomial models), allowing point interdependence, showed similar trends in the area surrounding the site to those displayed by Poisson regressions within the site. The linear model suggested an increase of intensity towards the (south)west, whereas the polynomial model strongly suggested an increase of intensity towards the south of the window (Fig. 6). Given the clustering nature of the Thomas algorithm, the random production of clusters and point scatters shows a more marked clustering process towards the (south)west (linear regression) or south (polynomial regression). Following the latter, the probabilities of finding clusters are substantially lower in any direction, except for the south of the excavated window. In contrast, the linear model allows a higher probability of finding clusters in any direction.

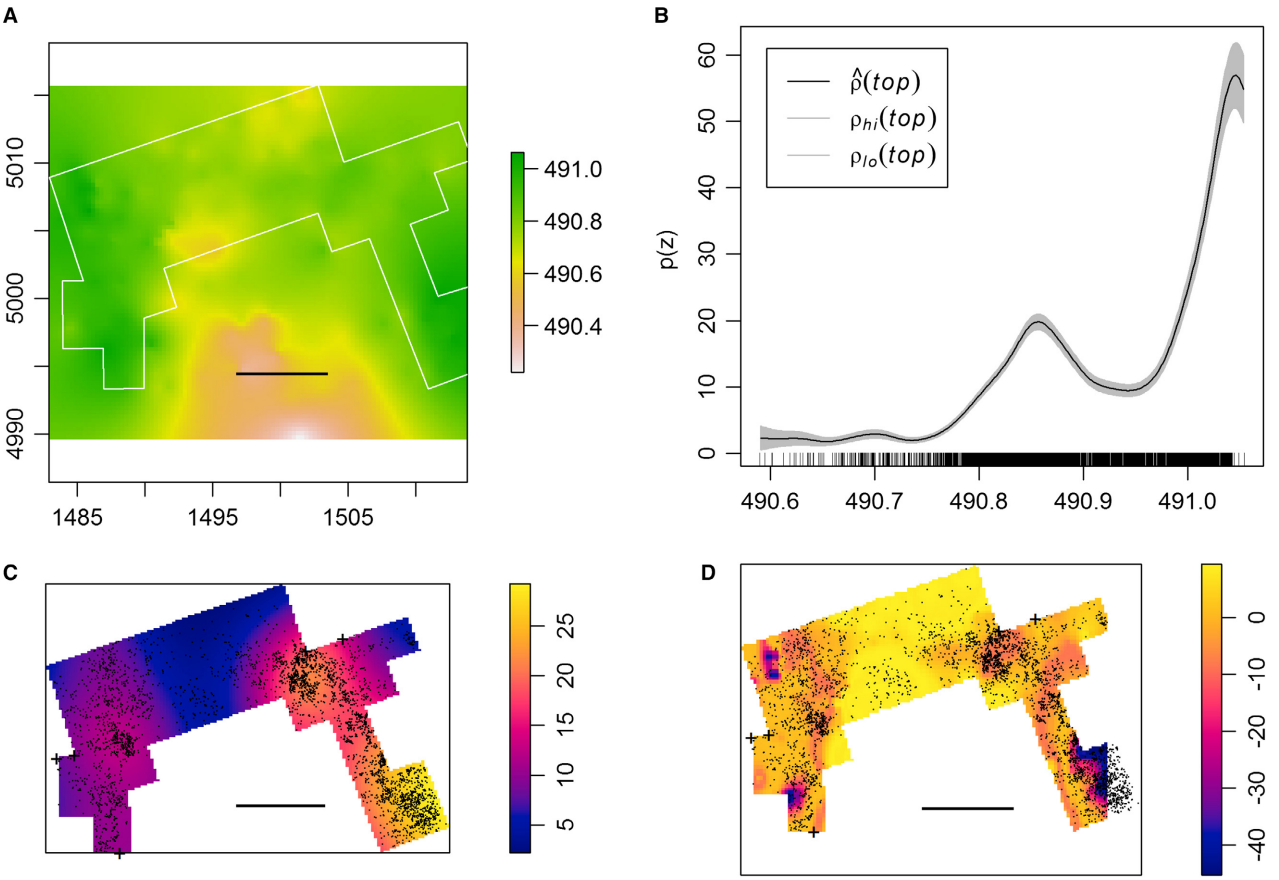

Fig. 8. A. Rasterized image of the topography of the DS palaeosurface. B. Estimation of parameter *p* (probability of density) based on depth (topography) using a resource selection function. C. Real density of excavated materials at DS. D. Validation test comparing the differences between predicted intensity and observed intensity. Scale bars = 6 m.

The more realistic simulation model generated by using random fields also emphasizes a south (mainly) and west (secondarily) increase of intensity (Fig. 7).

It should be stressed that these models provide an indication of the spatial trend of intensity and not an 'accurate' representation of the densities in the new 'unexcavated' window. The densest areas are expressed in relative and not absolute terms. The denser areas to the south in the polynomial models only indicate that the intensity trend is in that direction, even though the southern excavated area shows absolute values of materials that would match the predicted densities for the southernmost (unexcavated) area, whose absolute values are initially unknown and could eventually be as dense (or even absolutely less dense) as the excavated southern trenches. These models emphasize that a strategy targeting intensity should be used to prioritize intervention in the southern/southwestern sector of the unexcavated region. They also show that intensity should decrease towards the east and north.

It is interesting to note that the central area of the excavated window, which has been eroded, potentially contained large concentrations of materials, probably gaining intensity towards the south. It is unfortunate that this area, potentially as large as the excavated area, could have contained a rich archaeological record that has been lost. If so, the dimensions of the archaeological assemblage deposited on this palaeosurface have no equivalent and are substantially bigger than any early Pleistocene site known to date, especially when compared to sites dating to *c.* 2 Ma.

Once the spatial trend was modelled via these regression methods, we tried to understand the clustered inhomogeneous pattern by using the palaeosurface topography as the explanatory covariate. Both the Kolmogorov–Smirnov (D = 0.2851, p-value < 2.2e-16) and the Berman's tests (Z1 = 1.8842, p-value = 0.03953) suggested that the ground topography had a significant impact upon the distribution of materials (Fig. 8). This was independently confirmed by the smoothed kernel p estimator, which showed that below a depth of 490.8 m, there was no relevant density of materials and although this increased after that, it was on ground higher than 491 m that intensity spiked (Fig. 8B). The contrast between predicted and documented intensities showed a good match, with some discordances in the main cluster areas, which are much more dense in reality than predicted, as would be expected if the reason was behavioural and not just topographical (Fig. 8D).

A set of linear and polynomial regressions were used to create two different models of prediction and simulation of archaeological materials within the excavated window to determine the extent of the influence of the topography of the palaeosurface on the clustering and scattering of the point process (Fig. 9). The resulting models predicted the excavated data. The linear regression showed a better prediction of intensity (points per square metre) than the polynomial regression. The linear regression overemphasized the density of materials, whereas the polynomial regression showed a closer match in density of materials to the actual excavated assemblage. However, the polynomial regression failed at predicting the intensity of the clustering documented at the excavation, suggesting that the higher density of materials is attributed to behavioural and not just palaeotopographical attributes. Both regressions coincide in showing the low-density scatters to the east of the excavation, but the linear regression fails to reproduce the intensity of points documented to the north. Overall, the polynomial regression reproduced (proportionally) the excavated assemblage better. It showed the proportional distribution of materials very accurately compared to the excavated assemblage, but the density is lower than documented, as would be expected if density depended

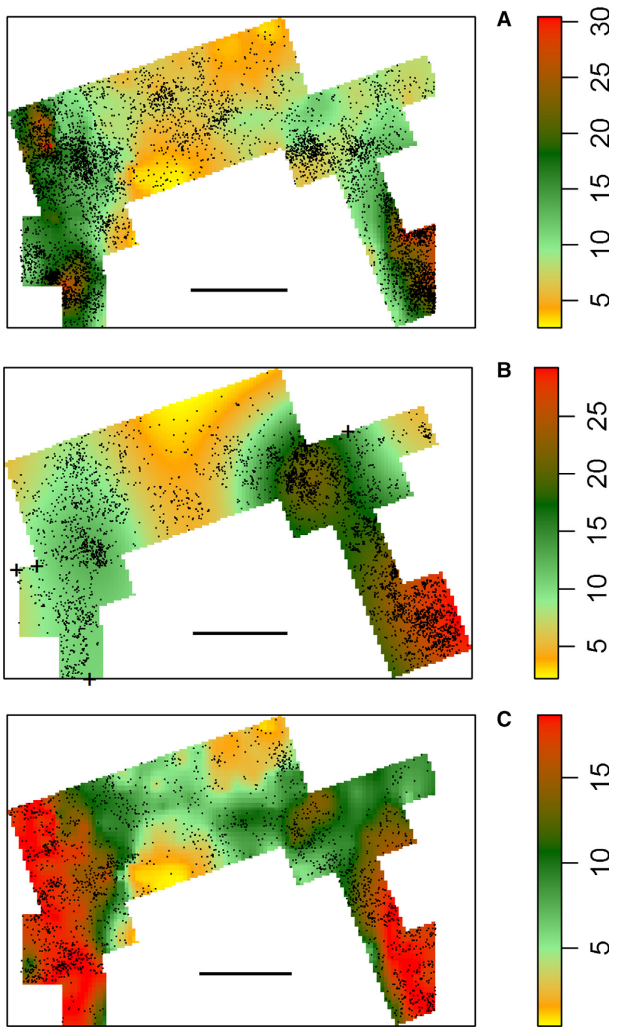

Fig. 9. A. Predicted density and simulated model of materials using a linear regression and a LGCP random field method. B. Documented density and distribution of materials at the excavation. C. Predicted density and simulated model of materials using a polynomial regression and a LGCP random field method. Scale bars = 6 m.

more on hominin behaviour than just the palaeosurface topography.

The control model provided a similar result to the dependent polynomial regressions. This enabled us to implement validation methods that showed that these regressions underestimate the density of points when using palaeotopography as a covariate. In Fig. 10, the top left panel represents the residual map with points representing positive residuals and the background colour scheme representing negative residuals (fitted intensity). The hotter colours of the smoothed residual field (Fig. 10, lower right) show positive residuals indicating an overabundance of points relative to the prediction of the model. The lurking variables (especially on the y coordinate; Fig. 10, top right and bottom left) also show that the model underestimates density in the southernmost clustering area and overestimates it on the northernmost area. Significant deviations are also documented on the x coordinate. The overall underestimation of intensity of the model further supports the behavioural nature of the point density and its general independence of the palaeotopography, despite its obvious influence.

Archaeological testing of the models

If considering the intensity documented in the excavated window, the three simulated models reproduce the documented density of materials in each area (Figs 4–6) with varying degrees of success. This allows predictions to test the efficiency of these models when extrapolating their predictive qualities to the area surrounding the excavated window. The predictions used here for the testing trenches were derived from the polynomial regression using the random field algorithm, given the characteristics of the assemblage (see above). Figure 7 shows four areas of interest adjacent to the excavated assemblage (a, b, c, d). These areas yielded different densities of remains once they were excavated (Fig. 11). The simulation pattern obtained through linear regression shows a moderate density in each of them (Fig. 5). The polynomial regression using a Thomas clustering method suggests that the intensity in the four areas is low (a, d), very high (b) and moderate (c), respectively (Fig. 6). The modification of this model introduced by a simulation based on random fields yields a similar pattern: low (a), high (b) moderate (c) and very low (d) (Fig. 7). The overall trend yielded by these simulation models suggests that the highest density of remains should be found in b, followed by c, a and d with the lowest density.

One could infer that the northernmost corner of the a area would contain a higher density of materials than modelled, as it is connecting one of the densest clusters excavated at the site thus far. This is why this spot was initially selected for maximizing the number of stone tool findings for phytolith and starch analysis. This explains why in Fig. 2 the forensic unit (ensuring decontamination from the exterior) was set up in that area right by the dense cluster of the adjacent trench. The predictive model indicated a paucity of materials in that area, which was subsequently confirmed through excavation.

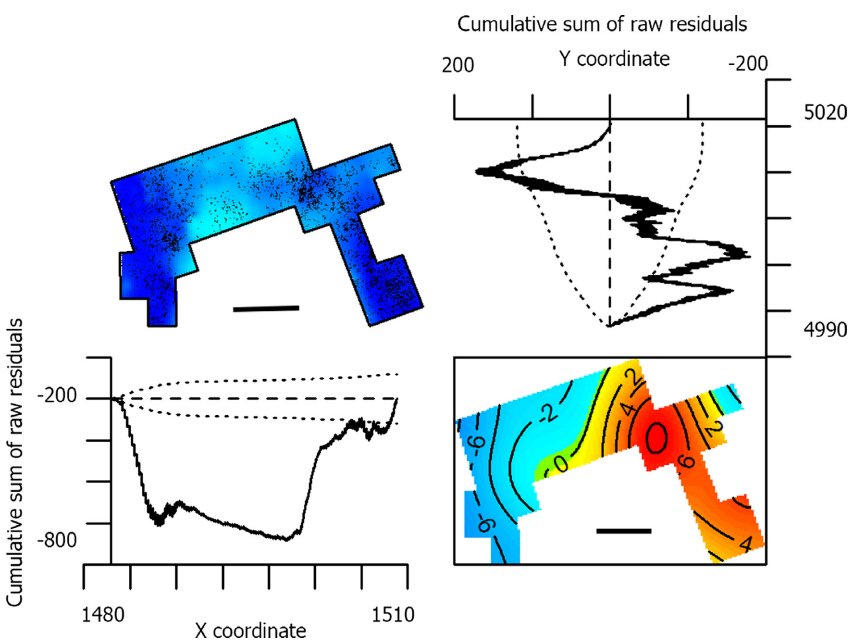

Fig. 10. Top left: residual estimation with points representing positive residuals and the background colour scheme representing negative residuals (fitted intensity). Bottom right: density map of residuals. Top right: lurking variable on the y coordinate. Bottom left: lurking variable on the x coordinate. See text for explanation. Scale bars = 6 m.

The correspondence of predicted and observed densities in the four trenches resulted in a Pearson correlation of 0.97 (p = 4.912e-08) and a perfect correlation (rho = 1, p = 0.083) using Spearman's method (Fig. 12).

Conclusions

The high degree of correspondence between the predicted densities and the observed densities at each of the four trenches confirms the great potential of spatial statistical analysis in archaeological research. This tight correlation between prediction and observation is probably caused by the interdependence of the components of the assemblage. This interdependence is probably caused by functional relationships between stone tools and bones and different types of bones amongst themselves. The interdependence of spatial objects, thus, creates spatial patterns that are not random and, therefore, that can be approached from multiple viewpoints: as independent spatial trends (using simple coordinates) or as dependent spatial trends (using other covariates such as topography, vegetation, chemical/isotopic signature, etc.). The fact that such a high correlation was found between prediction and observation in the present example does not mean that such a correlation will be found in other anthropogenic sites. We assume that there is a degree of variability, which will depend on the degree of interdependence and association of the materials that make up any given archaeological assemblage. Postdepositional disturbance can have a variable impact on these properties, thereby affecting the degree of matching between prediction and observation.

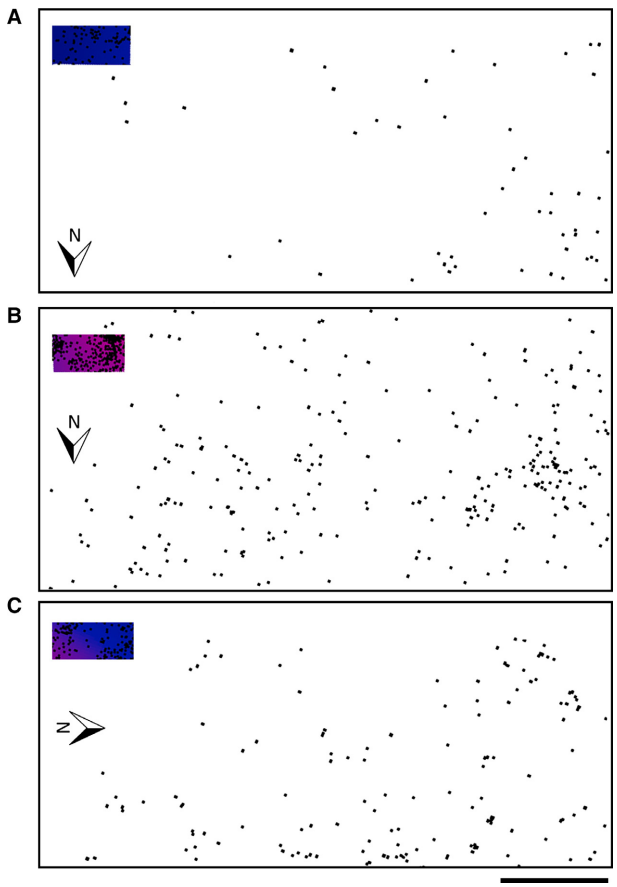

Fig. 11. Distribution of materials in the three testing trenches. Inserts (in colour) in each trench indicate predicted densities in each of them. Predictions suggested a low density in (A), high density in (B) and moderate density in (C). Scale bar = 1 m.

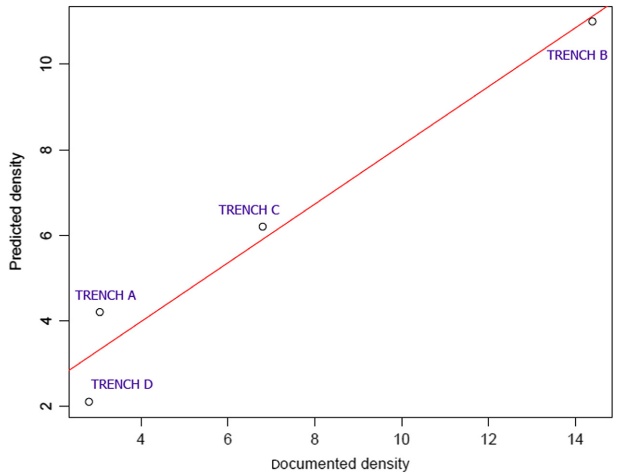

Fig. 12. Correlation between predicted and observed material densities at the four testing trenches.

If hominin behaviour creates spatial patterns, the analytical approach presented here can potentially contribute to the understanding of aspects beyond site formation and into the realm of hominin socio-economic behaviour at sites. On-going application of this approach to two anthropogenic sites in Olduvai Gorge (FLK Zinj and PTK) is yielding very informative results on the social behaviour displayed by hominins during carcass butchery and consumption (in progress). The baseline requirements are high, as only anthropogenic sites with minimal to moderate postdepositional disturbance (in which integrity and resolution are high) can be used. Alternatively, a substantial amount of actualistic/experimental work should be carried out within a spatial analytical framework to understand how different degrees of postdepositional disturbance affect the spatial properties that enable this type of modelling to be effective. In both cases, the possibilities that this type of analysis opens for the understanding of early human behaviour are unparalleled.

Acknowledgements. – We thank the Tanzanian Commission for Science and Technology (COSTECH), the Department of Antiquities and Ngorongoro Conservation Area Authority in the Ministry of Natural Resources and Tourism for permission to conduct research at Olduvai Gorge. We also thank the Spanish Ministry of Economy and Competitiveness for funding this research (HAR2013-45246-C3-1-P) and the Ministry of Culture for funding our research through their Archaeology Abroad Program. We are very thankful for the help

provided by Adrian Baddeley, Ege Rubak and Rolf Turner for some problematic coding. We thank them for creating the 'spastat' R library and for their excellent book. For their great work in the on-going field research conducted at DS, we would like to express our appreciation and acknowledgement to: Julius Sulley, Lazaro Sarwatt, Yacob Matle, Yona Thomas, Thomas Madangi, Nicolaus Dohho, Caroli Maole, Francis Fabiano, Sangau Letuma, Nicodemus Burra, Ibrahim Mathias and Shabany Bakari. We also thank J. Heaton, J. Rosell and R. Bivand for their constructive comments on an earlier draft of this paper. The authors have no conflict of interest to declare.

References

Arriaza, M. C. & Domínguez-Rodrigo, M. 2016: When felids and hominins ruled at Olduvai Gorge: a machine learning analysis of the skeletal profiles of the non-anthropogenic Bed I sites. *Quaternary Science Reviews 139*, 43–52.

Ashmore, W. 2002: "Decisions and Dispositions": socializing spatial archaeology. *American Anthropologist 104*, 1172–1183.

Ashmore, W. 2014: Practices of spatial discourse at Quelepa. *In* Braswell, G. (ed.): *The Ancient Maya and their Central American Neighbors*, 25–55. Routledge, London.

Baddeley, A. J. & Turner, R. 2005: Spatstat: an r package for analyzing spatial point patterns. *Journal of Statistical Software 12*, 1–42.

Baddeley, A., Rubak, E. & Turner, R. 2015: *Spatial Point Patterns: Methodology and Applications with R*. 848 pp. CRC Press, London.

Berman, M. 1986: Testing for spatial association between a point process and another stochastic process. *Applied Statistics 35*, 54–62.

Bivand, R. 2010: Exploratory spatial data analysis. *In* Fischer, M. & Getis, A. (eds.): *Handbook of Applied Spatial Analysis*, 219–254. Springer, Berlin.

Bivand, R., Pebesma, E. & Gómez-Rubio, V. 2013: *Applied Spatial Data Analysis with R*. Springer, New York.

Brughmans, T. 2012: Thinking through networks: a review of formal network methods in archaeology. *Journal of Archaeological Method and Theory 20*, 623–662.

Clark, P. J. & Evans, F. C. 1954: Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology 35*, 445–453.

Clarke, D. L. 1977: *Spatial Archaeology*. 397 pp. Academic Press, London.

Clarke, D. L. 1979: *Analytical Archaeologist: Collected Papers of David L. Clarke*. 554 pp. Academic Press, London.

Cressie, N. 2015: *Statistics for Spatial Data*. 928 pp. Wiley, New York.

Diggle, P. J. 2003: *Statistical Analysis of Spatial Point Patterns*. 349 pp. Arnold, London.

Diggle, P. J. 2013: *Spatial Analysis of Spatial and Spatio-temporal Point Patterns*. 300 pp. Chapman & Hall, London.

Donnelly, K. 1978: Simulations to determine the variance and edge-effect of total nearest neighbor distance. *In* Hodder, I. (ed.): *Simulation Studies in Archaeology*, 91–95. Cambridge University Press, Cambridge.

Dorman, M. 2014: *Learning R for Geospatial Analysis*. 360 pp. Packt Publishing Ltd, London.

Guan, Y. 2006: A composite likelihood approach in fitting spatial point process models. *Journal of the American Statistical Association 101*, 1502–1512.

Hodder, I. 1977a: Some new directions in the spatial analysis of archaeological data at the regional scale (macro). *In* Clarke, D. L. (ed.): *Spatial Archaeology*, 223–351. Academic Press, London.

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

Hodder, I. & Orton, C. 1976: *Spatial Analysis in Archaeology*. 270 pp. Cambridge University Press, Cambridge.

Leakey, M. D. 1971: *Olduvai Gorge: Volume 3, Excavations in Beds I and II, 1960–1963*. 463 pp. Cambridge University Press, Cambridge.

Møller, J. & Waagepetersen, R. 2003: *Statistical Inference and Simulation for Spatial Point Processes*. 292 pp. Chapman and Hall/CRC, Boca Raton.

Møller, J., Syversveen, A. & Waagepetersen, R. 1998: Log Gaussian Cox Processes. *Scandinavian Journal of Statistics 25*, 451–482.

Oyana, T. J. & Margai, F. 2015: *Spatial Analytics: Statistics, Visualization and Computational Methods*. 323 pp. CRC Press, Boca Raton.

Pebesma, E. J. 2004: Multivariable geostatistics in S: the gstat package. *Computational Geosciences 30*, 683–691.

Roger, S., Bivand, R. S. & Pebesma, E. J. 2013: *Applied Spatial Data Analysis with R*. 405 pp. Springer, Heidelberg.

Waagepetersen, R. 2007: An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics 63*, 252–258.

Wurzer, G., Kowarik, K. & Reschreiter, H. 2014: *Agent-based Modeling and Simulation in Archaeology*. 284 pp. Springer, New York.