



The strength of parthood ties. Modelling spatial units and fragmented objects with the TSAR method – Topological Study of Archaeological Refitting

Sébastien Plutniak

► To cite this version:

Sébastien Plutniak. The strength of parthood ties. Modelling spatial units and fragmented objects with the TSAR method – Topological Study of Archaeological Refitting. 2021. <hal-03274075>

HAL Id: hal-03274075

<https://hal.science/hal-03274075v1>

Preprint submitted on 29 Jun 2021

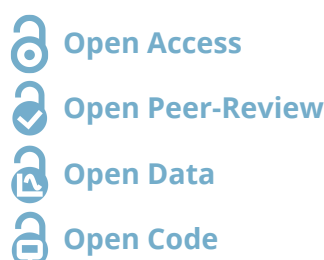
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

RESEARCH ARTICLE



Cite as: Plutniak, S. (2021) The strength of parthood ties. Modelling spatial units and fragmented objects with the TSAR method – Topological Study of Archaeological Refitting. OSF Preprints, q2e69, ver. 3 peer-reviewed and recommended by PCI Archaeo. doi: <https://osf.io/q2e69>

Posted: 24th June 2021

Recommender:
Hector A. Orengo

Reviewers:
Robert Bischoff, Matthew Peeples, and
an anonymous reviewer

Correspondence:
sebastien.plutniak@posteo.net

The strength of parthood ties. Modelling spatial units and fragmented objects with the TSAR method – Topological Study of Archaeological Refitting

Sébastien Plutniak¹

¹ Traces laboratory, CNRS, Université de Toulouse – Toulouse, France

This article has been peer-reviewed and recommended by
Peer Community In Archaeology
(<https://doi.org/10.24072/pci.archaeo.100010>)



Abstract

Refitting and conjoinable pieces have long been used in archaeology to assess the consistency of discrete spatial units, such as layers, and to evaluate disturbance and post-depositional processes. The majority of current methods, despite their differences, rely on the count and proportion of refits within and between spatial units. Little attention is paid to the distribution and topology of the relationships between fragments, although this is now known to have significant effects on archaeological interpretation. This paper presents a new methodological approach for refitting studies. The TSAR approach (Topological Study of Archaeological Refitting) draws on concepts and methods from graph theory to model the network of connections observed between refitting fragments. Measures of cohesion and admixture of spatial units are defined using the structural properties of the sets of refitting relationships. To ensure reproducibility and reusability, the TSAR method is implemented as an R package, which also includes a simulator generating refitting fragments scattered in two spatial units. The advantages of the topological approach are discussed by comparing it to: 1) the results of a survey in which archaeologists were asked to rank examples of stratigraphic admixture; and 2) other computational methods. The approach is applied to simulated data, and empirical data from the Liang Abu rock shelter (East Borneo) are presented. Finally, the use of the TSAR simulation approach to test different scenarios of site formation processes is demonstrated.

Keywords: Refitting; Graph analysis; Network analysis; Stratigraphy; Post-depositional process; Taphonomy; Software

1 Introduction

Modern studies of archaeological stratigraphy involve numerous specialists, with each specialist shedding light on the relevant distinctions in an archaeological sequence. Direct observation during excavation, geoarchaeology, sedimentology and pedology, chronometric results, the spatial study of remains, technological and stylistic analysis of artefacts: all these complementary approaches, convergent or divergent, contribute to identifying significant changes in the vertical sequence and, consequently, through time (Lyman and O'Brien, 1999). In this context, “refitting” of archaeological remains, namely the identification of fragments that were parts of the same initial object, has long been used as a method to assess the integrity of archaeological discrete spatial units. Inter-layer displacement of remains is used as evidence of post-depositional disturbances (Myers, 1958, Villa, 1982, Hofman, 1986, Ziesaire, 1990, Barthès, 1994, Bordes, 2000, Morin et al., 2005), intra-site refitting is interpreted as evidence of site re-use (Cahen and Moeyersons, 1977), and inter-sites admixture is considered



as evidence of the movement of people (Schaller-Åhrberg, 1990).

In the last three decades, refitting analysis has become an extensively studied field, addressed through several conferences and collective publications. This field has generated numerous methodological improvements, which cannot be summarised here (Cziesla et al., 1990, Hofman and Enloe, 1992, Schurmans and De Bie, 2007; for a history, see Schurmans, 2007). Post-depositional processes such as trampling (Villa and Courtin, 1983) and movements have long been recognised as possible disturbance processes (Wood and Johnson, 1982, Johnson, 1989). In this context, about thirty years ago, Hofman urged archaeologists to not simply use refitting to acknowledge that disturbances had occurred, but rather to document “the degree of such movements”, namely to quantify them (Hofman, 1992, p. 4). In this regard, many methods have been proposed to quantify refitting in always more refined ways, for example by distinguishing between three types of lithic refits (Cziesla, 1990), or between six types of pottery refits (Bollong, 1994), or even between seven types of lithic refits (López-Ortega et al., 2011).

However, despite the increasing refinements of these methods, all were based on the *count* and the proportion of internal and external refits. The main argument of this paper is that quantifying the number of relationships alone can be misleading and should be complemented by quantifying the *topology* of these relationships, i.e. the “structure” of the network of refitting relationships between pieces. Note that paying attention to the topology is not a completely new idea in this context (Cziesla, 1990, Michel, 2002). However, previous attempts were limited to manual graphic procedures, based on the count of (different types of) relationships, and were not generalised and implemented into data processing procedures. There are good reasons to integrate the topology of refitting relationships, as illustrated by the theoretical examples in Figure 1: although the refit count and the proportion of internal versus external relationships are equal, conclusions on the validity of the layers and the initial location of the remains are different. To overcome these ambiguities, we use graph theory to model sets of refitting relationships and study their topology. This paper aims to: 1) define and implement the Topological Study of Archaeological Refitting (TSAR), a method to quantify the “strength” of archaeological units (e.g., layers), based on the topology of parthood relationships of the fragments, i.e., relationships between parts and wholes; 2) compare TSAR with previous and alternative methods; and 3) present results from simulated data and an application for field archaeological data.

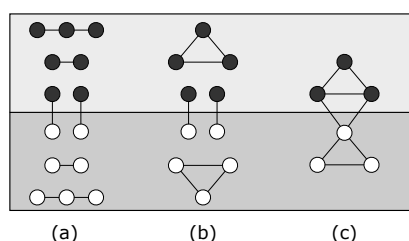


Figure 1. The need to consider topology: three examples of two layer with internal refitting ($n=6$) and inter-layers refitting ($n=2$). Although the numbers of relationships are equal in all examples, their archaeological interpretation are very different. The distinction between the two layers would be considered as relevant in (a); relevant with higher confidence about the objects' initial location in (b); doubtful in (c).

2 Material and methods

2.1 Preliminary definitions

2.1.1 Layers and spatial units

The problem addressed by the topological method (TSAR) initially relates to stratigraphical analysis (Harris, 1979, Lyman and O'Brien, 1999, May, 2020). Archaeological observation starts by distinguishing spatially ordered volumes, called "layers" (or "stratum", "levels", or "archaeological units"). In the context of this paper, layers will be considered. However, layers are only one example of archaeological discrete spatial unit¹, characterised by their emphasis on vertical orientation and related by binary relationships (such as *being above*, *being below*, and *being inside*). In this paper, it has to be kept in mind that the topological method, used to assess the reliability of the limit of the layers and the post-depositional processes, can be applied to any type of archaeological spatial unit as long as it contains fragmented objects (e.g., between the inside and outside of holes, buildings, specialised areas).

2.1.2 Connection and similarity relationships

Connection Consider a material object fragmented into multiple pieces. All fragments share the same parthood relationship with the initial object, of which they are parts. More precisely, the observed fragments *are* disjoint parts of the initial object, and *were* connected parts of this object. In the TSAR approach, the term "connection" is used as a shorthand to refer to the connection relationship which *existed* in the past, before the object was broken into fragments (Figure 2). Archaeologists deduce connection relationships from the symmetry and the possibility of large *contact* between substantial areas of the surfaces of two fragments,

¹On archaeological units, see O'Brien and Lyman, 2002.

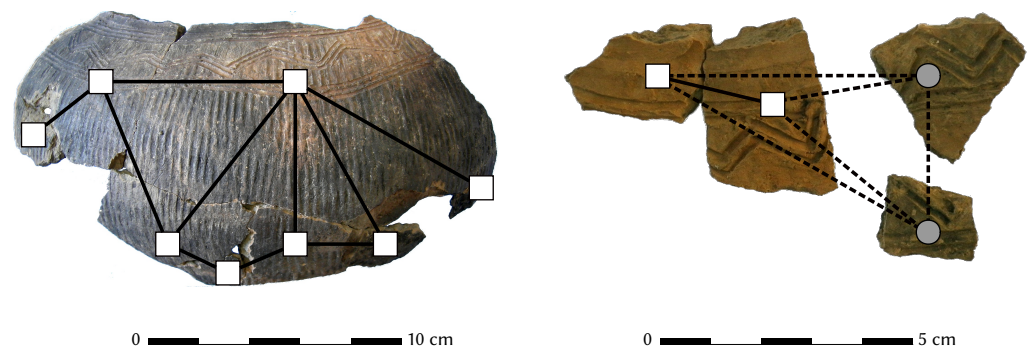


Figure 2. Pottery sherds from Liang Abu related by connection relationships (solid line between white squares) and similarity relationships (dashed lines between grey circles).

which can be physically adjusted (the fragments “refit”). The connection relationship is the fundamental relationship in the TSAR method.

Similarity A second type of relationship is distinguished, namely “similarity”². Similarity relationships exist between fragments considered as sharing enough common features (motif, clay, inclusions, etc.) to state they are (and were) parts of the same initial object.

2.1.3 Topological properties

Contiguity Determining contact between fragments raises a practical difficulty because it is also related to a classic issue regarding contiguity. Inferring a connection relationship between two fragments both with very small surfaces is ambiguous; thus two types of contiguity are distinguished in spatial analysis, namely “Queen” and “Rook” contiguity (Figure 3)³. Queen contiguity refers to contiguity determined by the sharing of at least a point (or a line in 3 dimensions), whereas Rook contiguity is determined by the sharing of a line (a surface in 3 dimensions). When recording data, the choice of the type of contiguity is crucial because it has significant effects on the numerical values used in the computations.

In the context of archaeological fragments analysis, Rook contiguity is more relevant for practical, reliability, and conceptual reasons. At the daily-life, meso scale of analysis used in refitting analysis, a “point-like” or “line-like” contact is not enough to infer with certainty a past connection between two fragments. Consequently, Rook contiguity must be favoured since surfaces have a major role in this concept.

²In a previous and slightly different attempt to determine archaeological structures, Moberg distinguished two relationships (connection and inclusion) which can form two types of clusters, on the similarity and proximity relationships, respectively (Moberg, 1971, pp. 554–555).

³Which are also known as Moore and Von Neumann neighbourhood, respectively (Gray, 2003).

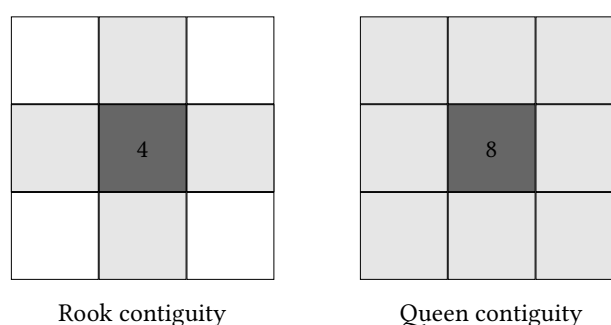


Figure 3. “Rook” and “Queen” contiguity. The labels indicate the number of contiguous neighbours of the cells. Note that the difference between the two series of values is not proportional.

Graph planarity The concepts and tools of graph theory can be used to model the connection relationships of archaeological fragments. This implies defining the type of graph required, e.g., using undirected edges (since a connection relationship is symmetric). In the context of the relationships between (parts of) material entities, it is also relevant to determine whether the graphs must be planar or not. In graph theory, a graph is planar if all intersections of its edges in the diagram are vertices of this graph (no edges cross each other) (Ore, 1962, p. 6). An exploratory analysis identified that some specific sets of connecting archaeological fragments have to be represented by non-planar graphs. This would be rather rare in the case of pottery sherds (see Figure 4) but more frequent for lithic artefacts (e.g., Sisk and Shea, 2008, Fig. 5, p. 491)⁴. Accordingly, the archaeological context and, consequently, the nature of the objects, the number of fragments conserved, and the number of connection relationships identified will all determine the likelihood to require or not non-planar graphs (e.g., Paleolithic stone artefacts, sparse Neolithic simple-shaped potteries, numerous medieval urban vessels with complex shapes).

2.2 Measures

2.2.1 The TSAR approach: topological study of archaeological refitting

Taking topology into account is the main principle of the TSAR method. However, related conceptual choices are clarified below.

Types of relationships Two types of relationships between fragments were distinguished above: connection and similarity. The TSAR method focuses on connection relationships, for two reasons. First, similarity relationships imply that degrees of certainty are accounted for,

⁴Note that these examples differ regarding the status of intentionality in the fragmentation process: breaking is generally unintentional for pottery, whereas for lithic objects it is inherent to the production of the object by humans.

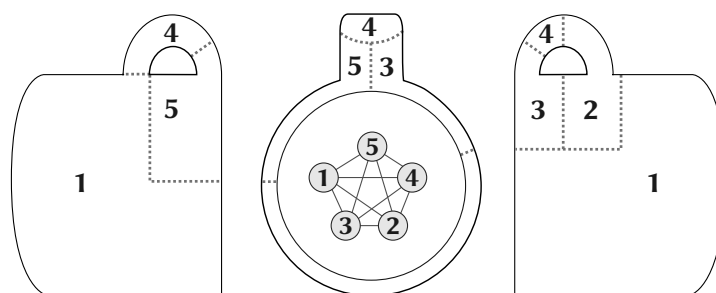


Figure 4. Non-planar fragmentation graph. Virtual example of a pot with a handle fragmented into five pieces. The corresponding non-planar graph is represented in the middle of the figure.

whereas connection is dichotomous (one can state if there is or not a refitting). Addressing degrees of certainty makes the mathematics much more complex, as does the second reason: similarity is a transitive relation (A is similar to B, B to C, so C to A) whereas connection is not.

Consideration of sample size The quantity of archaeological remains found in different layers is rarely equal. This empirical constraint must be addressed in the design of the method. In archaeological reasoning, it is acceptable to assume that the determination of a layer is more liable if it contains more material. However, the TSAR method only concerns remains implied in connection relationships. Fragments without connections, “singletons”, are excluded. Therefore, the interpretation of the result considering the total quantity of remains is left to the archaeologist. Methodological coherence justifies this choice. This is grounded on the principle that an object or fragment had a single location before the fragmentation and dispersion processes occurred. Fragmentation analysis is about observable archaeological evidence of these processes. Singletons are weak proof of their initial location because they are only related to their layers (by a relationship of inclusion). On the contrary, the more fragments sharing a similar location are connected, the more their association with their observable inclusion within a given layer is strengthened. Consequently, the difference in sample size is addressed by determining the size of a layer from its number of connected fragments and connection relationships. This approach distinguishes between a case where 10 connection relationships are observed between a layer containing 20 fragments and a layer containing 80 fragments, and a case where 10 connection relationships are observed between two layers containing 50 fragments each.

Admixture and reliability The first aim of the TSAR method is to evaluate the reliability of a distinction between two archaeological spatial units and to interpret the site formation process. Achieving this implies evaluating and comparing the degree of consistency of the units, how they are “self-adherent” to themselves, their *cohesion*.



In the TSAR method, the opposite concept of cohesion is *admixture*. Cohesion is a property of a spatial unit whereas admixture is a property of a pair of spatial units. The concept of admixture supposes, first, to distinguish two containers and their respective content (e.g., layers and fragments) and, second, to determine that a part of the content is not within its expected container (e.g., fragments which moved to a different layer). Considering two spatial units, the TSAR method proceeds by defining a virtual third unit including the fragment and connection relationships at the intersection of the two units. Therefore, the admixture corresponds to the part of the global cohesion which is not specific to one of the two units under study.

Cohesion and admixture measurements are used to 1) quantify the reliability of a distinction between two units; and 2) suggest hypotheses about the post-depositional processes which might have disturbed the units. Comparing the results of the two measurements generates four possible situations (Table 1 and Figure 5). Observing two layers with similar cohesion values and low admixture validates the distinction between these layers and explains their admixture by the inter-layer movement of fragments (case 4 in Table 1). The case of two layers with very different cohesion values but low admixture validates only one layer and suggests transport of fragments from this layer to another (case 3). On the contrary, the distinction between two layers would be questionable for layers with similar cohesion values and high admixture (case 2). Finally, for layers with high admixture and very different cohesion values, the uncertainty is too high to validate any of the layers (case 1).


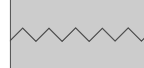
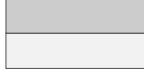

Given that the past series of the states of the spatial units are (by definition) unknown, and that there is no general law for site formation, post-depositional processes cannot be directly inferred from the cohesion and admixture measurements of the archaeological units. However, based on these measurements and general knowledge about the archaeological context, different hypotheses about the initial state of the units can be formulated and tested using the TSAR simulator, overcoming the impossibility of direct observation of the past states of the archaeological deposit.

2.2.2 Edge weighting

Relationships receive particular attention in the TSAR method since connection relationships between fragments are central in this approach. Using graph modelling, relationships are represented by edges, which are weighted to reflect their significance in this archaeological context. An edge weight will be higher if it connects fragments that are: 1) connected to many other fragments; 2) parts of triangles (A is connected to B, B to C, C to A); and 3) located in a spatial unit containing many other fragments and connections.

Concepts and methods from graph theory are used to model and achieve this principle. Considering two layers, three sub-graphs are extracted, including respectively the edges within layer 1, within layer 2, and between layers 1 and 2. In each sub-graph, the edges are weighted with the sum of the degree of the nodes they connect, modified by a “structural” factor (based

Table 1. Interpretation of the cohesion and admixture values. The results characterise the reliability of the distinction between two archaeological units and also suggest hypotheses about their formation process: 1) general uncertainty; 2) transport of fragments within one initial unit; 3) transport from one certain unit; 4) transport between two certain units. See also the practical examples of actual fragmentation graphs (Figure 5) and simulated graphs (Figure 10).

Admixture	Difference between the cohesion values	
	high	low
high	1. 	2. 
low	3. 	4. 

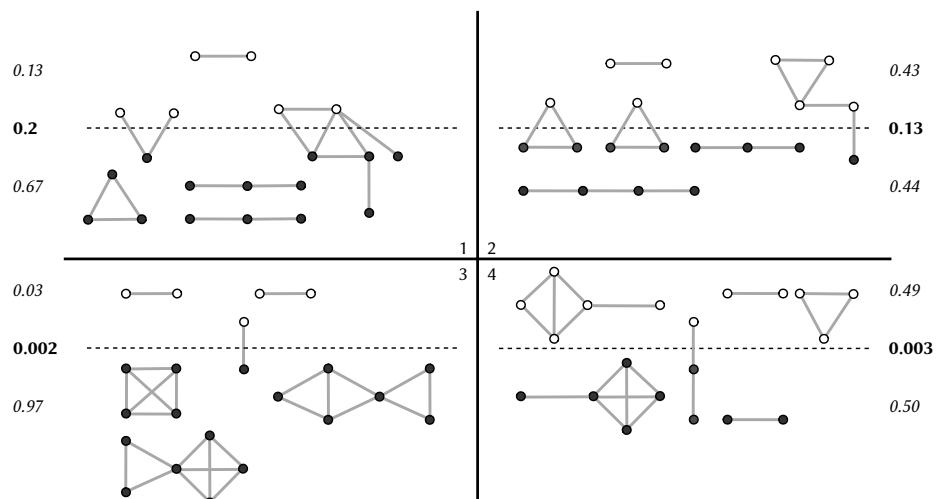


Figure 5. Illustrations of the four possible interpretations presented in Table 1. The graphs were generated using the TSAR simulator. Their cohesion values are in italics and their admixture values in bold font.



on the local transitivity of the vertices connected by the edge⁵) and a “size” factor (based on the number of connected fragments and connection relationships in the sub-graph)⁶:

$$W(E_{ij}) = (d_i + d_j) \times \left(3 - \frac{2}{1 + (tr_i + tr_j)/2} \right) \times \left(1 - \frac{1}{\sqrt{(V_{sub} + E_{sub})}} \right)^2$$

with d_i and d_j the degrees of the vertices i and j , tr_i and tr_j their local transitivity values, V_{sub} the number of vertices in the sub-graph and E_{sub} the number of edges in the sub-graph.

2.2.3 Cohesion

The cohesion of a layer is determined from the number of fragments and connection relationships it contains and from the strength of these connections (represented by the edge weights). Cohesion is always determined in the context of a comparison between two spatial units. This constraint is justified by: 1) the need to reduce the analysis to a simple question, the pair of spatial units being an elementary scale of analysis, which corresponds to: 2) the fact that, in stratigraphical analysis in archaeology, spatial units are related by basic binary relationships (*above to*, *below to*, *included in*, etc.). (This does not prevent us from studying more than two spatial units by repeating the procedure for each related pair of units.) The cohesion value of a spatial unit is given by:

$$\text{cohesion}(\text{unit}) = \frac{\sum W_{unit} + V_{unit}}{\sum W + V}$$

with V_{unit} the number of vertices in the unit, V the total number of fragments in the two units considered, W_{unit} the sum of the edge weights within the unit and W the sum of all the edge weights. Note that the relative size of each unit is also included in this computation. Considering a pair of units, their respective cohesion values range between 0 and 1, with values towards 1 reflecting higher cohesion and values towards 0 for lower cohesion. The sum of the cohesion values of a pair of spatial units is always equal to 1. For example, two layers with no inter-layer connections and containing the same number of fragments and the same patterns of connection relationships will have cohesion values equal to 0.5.

2.2.4 Admixture

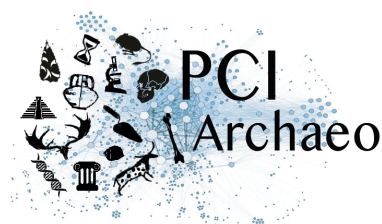
The admixture value of a pair of spatial units is equal to the “cohesion” value of a virtual third unit containing the fragments and relationships at the intersection of the two units. Therefore, it is simply computed as:

$$\text{admixture}(\text{unit1}, \text{unit2}) = 1 - (\text{cohesion}_{\text{unit1}} + \text{cohesion}_{\text{unit2}})$$

Results range from 0 to 1, with 0 for unmixed units and values towards 1 for very mixed units.

⁵Transitivity is also called the clustering coefficient, see Wasserman and Faust, 1994, p. 243.

⁶Unconnected vertices are not considered. See supplementary material, section 4.2.2 “Relative sizes of the components”.



2.2.5 Alternative methods to compare

In this section, the computation method used in TSAR is compared to three alternative methods.

Edge count The simplest approach, commonly used by archaeologists, is the ratio of the number of relationships between two different layers over the total number of relationships within the layers. The limitations of this method were presented in introduction.

Modularity Modularity is a “quality” measure for graph partition (Newman, 2006, Clauset et al., 2004). Given a graph with two groups of nodes, its modularity is the fraction of edges that fall within group 1 or 2, minus the expected number of edges within groups 1 and 2 for a random graph with the same node degree distribution as the graph under study. Many methods to detect “communities” in graphs (classes of nodes with dense relations) work by optimising modularity.

In the context of archaeological analysis, the association (inclusion) between fragments (nodes) and layers (node attribute) can be modelled as a partition of the nodes, and therefore modularity might appear as a relevant method to evaluate communities. However, there are two reasons against this claim. First, modularity is known to have a low sensitivity for small node groups, a problem called “resolution limit” (Fortunato and Barthélemy, 2007); archaeological graphs are often small. Second, modularity assumes that potentially all nodes can be connected. This goes against one of the properties of archaeological fragmentation graphs determined by the ontology of material objects: fragments from different initial objects cannot be connected and, in addition, fragments from the same object but located in non-adjacent positions in this object cannot be connected. Modularity must therefore be abandoned in the context of archaeological fragmentation analysis. It will nevertheless be included in the analysis for comparison purposes.

Topological admixture (edge betweenness centrality) The topological admixture, presented above, can be modified by using edge betweenness centrality to weight the edges. This variant has the advantage of relying on a previously defined and well-known metric. In a graph, the edge betweenness centrality of an edge is defined by the number of shortest paths going through this edge (Girvan and Newman, 2002). However, our exploratory analyses showed that the basic behaviour of the edge betweenness gives results inverse to our premise for archaeological interpretation (i.e., a connection between fragments that are densely connected to the fragments in their layer must be given a low weight)⁷. This can be overcome by adjusting the method without, but it does not resolve the second issue, namely assigning 0 to connections between two single fragments. However, this is a non-trivial issue, given that pairs of connecting isolated fragments are frequently observed in archaeology.

⁷See supplementary material, section 4.1.1 “Application of the four methods”.

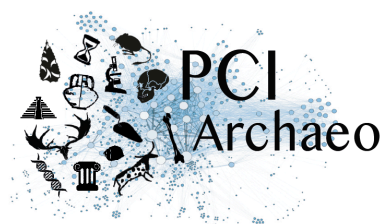


Table 2. Parameters of the TSAR simulator.

Parameter	Type	Description
components	integer	number of initial objects
vertices	integer	number of fragments
edges	integer	number of connection relationships
balance	numerical]0;1[proportion of fragments in each layer, before post-depositional processes
components balance	numerical]0;1[proportion of components in each layer
disturbance	numerical [0;1]	proportion of fragments likely to move from a layer to the other one
asymmetric transport	integer [1;2]	whether to only disturb the fragments from layer 1 or layer 2
aggregation factor	numerical [0;1]	when applying fragmentation, increase the likelihood for objects with more fragments being selected
initial layers	integer [1;2]	number of initial hypothetical layers
planar	Boolean	whether generating only planar graphs

2.3 The TSAR simulator

To test the different methods and to generate data to compare with empirical observations, a simulator was designed, implemented in R language (R Core Team, 2020) and included in the *archeofrag* package (Plutniak, 2021). It can be set with several parameters (Table 2).

2.3.1 Algorithm

The TSAR simulator implements a model of archaeological spatial fragmentation, focusing on the topology of the relationships between fragments. Fragments can be located in two different spatial units. An initial object is first broken into two fragments (Figure 6). The second fragmentation is then applied to one of the two fragments. At this stage, four different results are possible. For 10 iterations, the number of possibilities has an order of magnitude of 10^9 . However, as illustrated in Figure 6, the number of different graphs to model these possibilities is always lower, since the relative connections between the fragments are considered regardless of their orientation in space (*right of*, *above of*, etc. are not considered). No assumptions are made about the probability of an object being fragmented or transported. These parameters are intended to be deduced from empirical observation, chosen by the user, or simulated using multiple values⁸.

The TSAR simulator implements this approach, using the algorithm 1, which can be sum-

⁸Different probabilities that a sample move to a different layer were used in previous models of post-depositional mixing, occasionally using different values for above or below, adjacent or non-adjacent, close or distant, layers (Rowlett and Robbins, 1982, p. 79, Brantingham et al., 2007, Caron et al., 2011).



marised as:

1. generate one or two initial spatial units
2. disseminate n initial objects into the spatial unit(s)
3. select an object or a fragment
4. break it into two fragments
5. return to [1] while the number of fragments and/or the number of connection relationships is not reached.

When the simulator is set with two initial spatial units, steps 2–5 are executed separately for each unit before integrating the resulting graphs into a single graph (as illustrated in Figure 7).

```

Input :  $n, v', e'$ , with
          $n$ , the initial number of vertices of  $G$ ;
          $v'$ , the final number of vertices of  $G$ ;
          $e'$ , the final number of edges of  $G$ ;
Result:  $G(V, E)$  a graph with  $|V| = v'$  vertices and  $|E| = e'$  edges
begin
   $V \leftarrow \{v_1, \dots, v_n\}$  // initialise with  $n$  vertices
   $E \leftarrow \{\}$  // initialise with 0 edges
  while  $|V| < v'$  or  $|E| < e'$  do
     $x \leftarrow \text{sample}(V)$  //  $v_i \in V$  with  $i$  randomly generated in  $[1, |V|]$ 
     $V \leftarrow V \cup \{y\}$  //  $y$  is a new vertex:  $y = v_{|V|+1}$ 
     $E \leftarrow E \cup \{(x, y)\}$ 
  end
  return  $G(V, E)$ 
end

```

Algorithm 1: Algorithm of the TSAR simulator (when set with 1 spatial unit as initial condition).

2.3.2 Validation of the simulator

Multiple tests were run to validate the simulator by comparing input parameters and the properties of the graphs generated⁹. In summary, results were 100% accurate for the number of objects, fragments, and connection relationships. However, since these numbers must be integers, rounding of numbers was applied, leading to slight inaccuracies for the balance (median inaccuracy = -0.04 ± 0.08 on a scale of 0 to 1) and disturbance (median inaccuracy = -0.07 ± 0.08) parameters. In addition, the effect of the “aggregation factor” can vary since it is based on random selection. Finally, the simulator has acceptable support for scaling (decreasing the size of the graphs does not affect the other parameters) and the planarity constraint has no side effects.

⁹See supplementary material, section 4.2 “Evaluation of the TSAR simulator performances”.

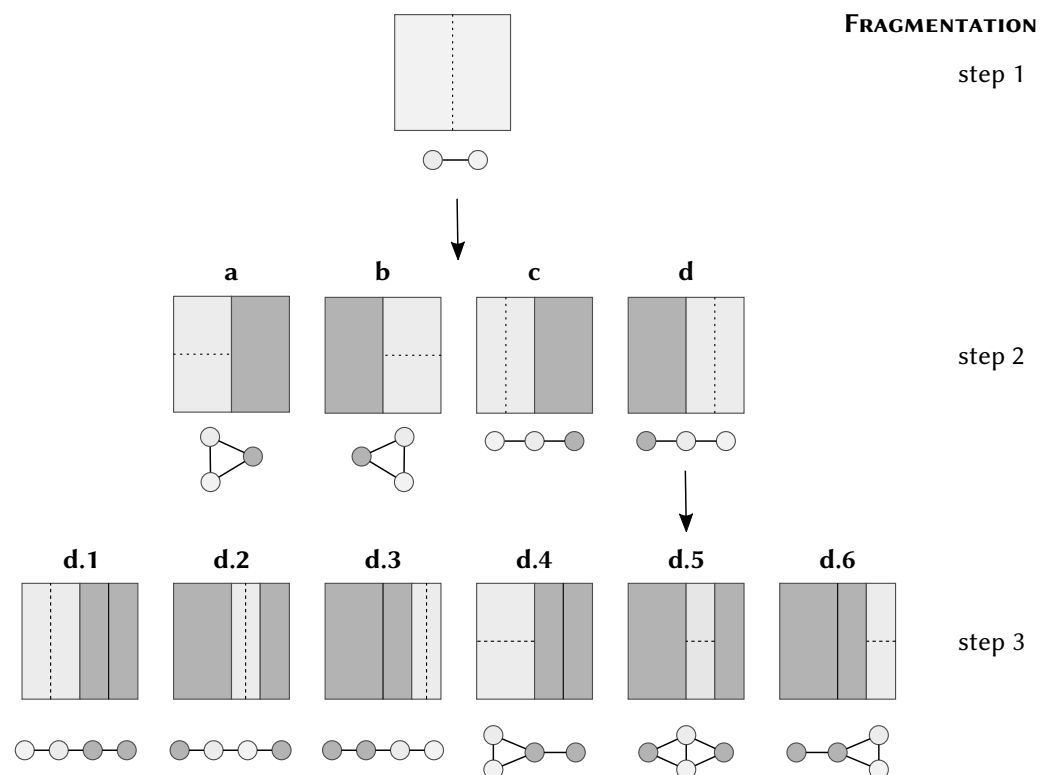


Figure 6. Model of archaeological fragmentation. Unchanged fragments are coloured in dark grey. The border between newly created spatial units are represented by dashed lines. In step (1) an object is fragmented into two fragments. Four different configurations are possible in step (2) (a, b, c, d). In step (3), each configuration gives six other possibilities, 24 in total (only those related to configuration (d) are represented here).

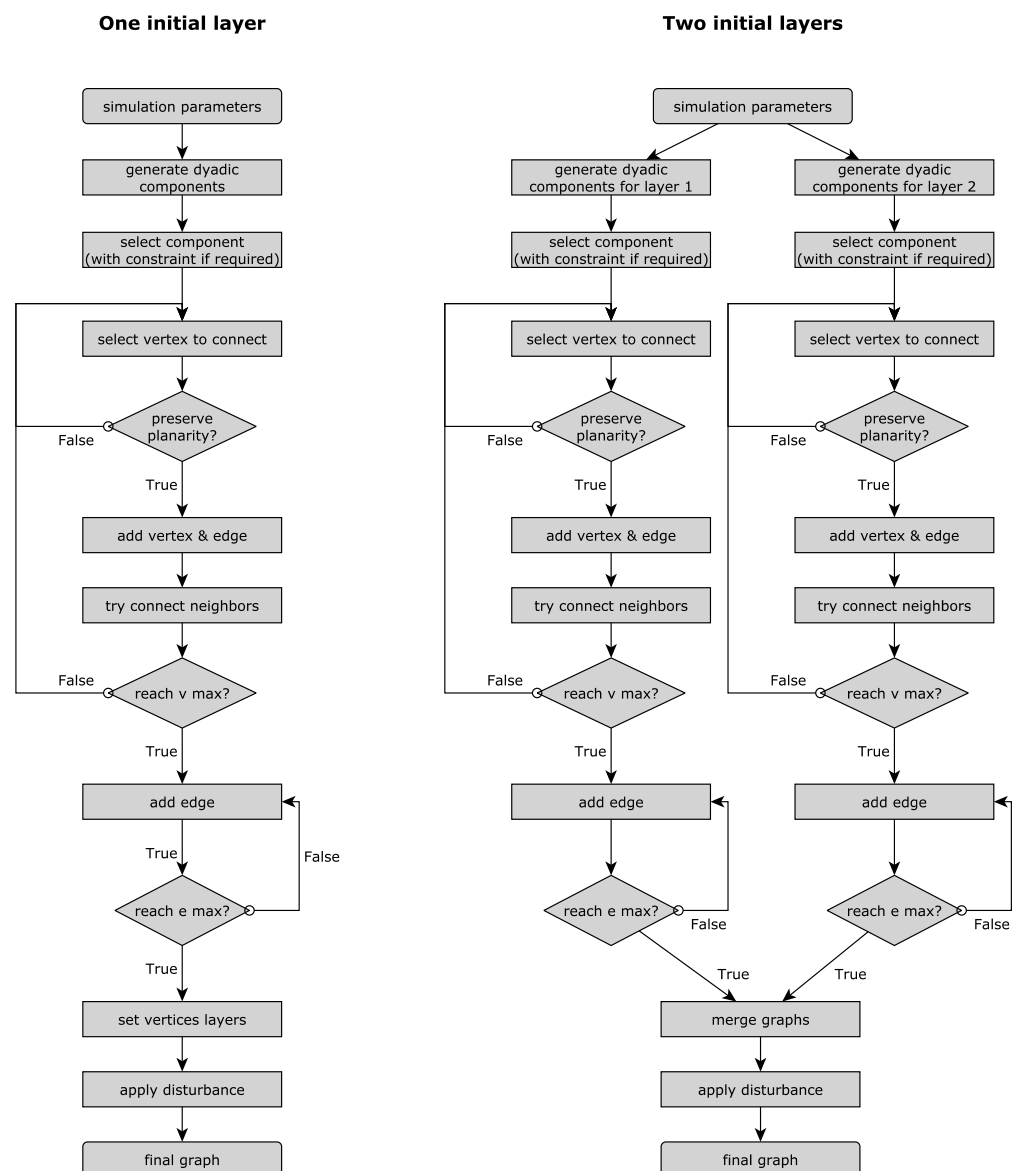


Figure 7. Flowchart of the TSAR simulation function for the two possible initial conditions: one spatial unit (left) and two spatial units (right).

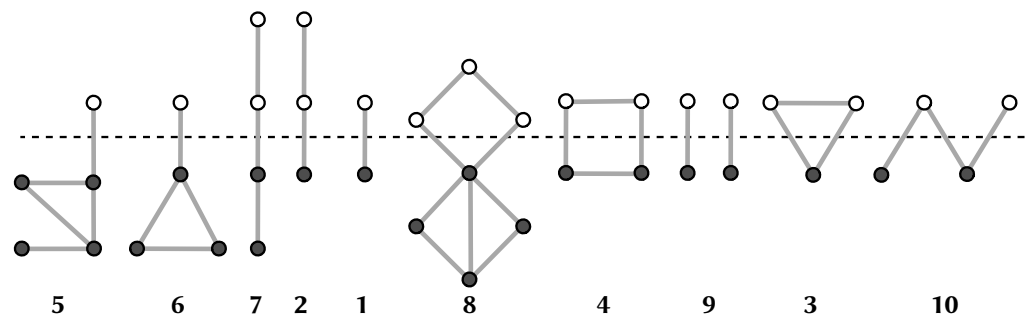


Figure 8. Theoretical examples. The colour of the nodes gives the spatial unit where the fragment was located. The graphs are ranked by topological admixture value.

2.4 Data

Three datasets were used in this study: a set of theoretical small fragmentation graphs, the connection relationships between pottery sherds found in the Liang Abu rock shelter, and data generated with the TSAR simulator.

2.4.1 Theoretical examples

A set of ten simplistic fragmentation graphs was defined, in which connected fragments are located in two spatial units (Figure 8). These examples were used for two purposes: to test and illustrate the different methods, and to collect and reflect archaeologists' intuitive estimations of admixture between the two units in each case. Using consensus modelling¹⁰, archaeologists were asked to rank graphs 1–8 from the case where the layers are most distinguished (least mixed) to the case where they are least distinct (most mixed)¹¹: 30 archaeologists were surveyed, generating 35 ranks in total, corresponding to 30 different solutions.

2.4.2 Liang Abu dataset

The real-world dataset used in this study comes from Liang Abu, a rock shelter located in East Borneo and excavated in 2009 and 2012 (Plutniak et al., 2016, see also Figure 2). Pottery was found on the surface, in layer 1 and 2, raising issues specific to shallowly buried sites (Surovell et al., 2005). Two ¹⁴C datings on charcoal from layer 2 gave reliable results at 1672±21 uncal. BP and 1524±22 uncal. BP, providing a *terminus post quem* for pottery. All sherds show similar stylistic and morphological features but layers 1 and 2 have different sediments; layer 1 is yellowish silt sediment and layer 2 is a gravel line mixed with dark brown silt sediment. This raises an interesting stratigraphical problem: is the distinction between layer 1 and 2 reliable

¹⁰For a review of the literature on the evaluation of classification consistency, and a similar approach applied to pottery typology, see Whittaker et al., 1998, pp. 138–143; for a study of variation in stone artefacts recording by 15 observers, see Gnaden and Holdaway, 2000.

¹¹Graphs 9 and 10 were included after the survey.



Table 3. Liang Abu: number of fragments, connection relationships, and maximal number of initial objects by layer (numbered from 0 to 2).

	Fragments	Connections	Objects
all layers	78	56	30
layers 0 & 1	29	22	11
layers 1 & 2	72	52	28

Table 4. Liang Abu: distribution of the connection relationships within and between the three layers.

	0	1	2
0	4		
1	0	18	
2	0	3	31

and how to interpret this distinction in the formation process of the site? The study of the connection relationships between fragments contributes to answering these questions. The data are summarised in Table 3 and Table 4.

Applying the TSAR method includes four steps: 1) weighting the connection relationships; 2) measuring the cohesion and admixture of the spatial units; 3) interpreting the reliability of the distinction between the units and the possible disturbances, based on these measures; and 4) optionally, testing this interpretation using simulation.

3 Results and discussion

Results are presented in three sections: 1) demonstrating the relevance and reliability of the TSAR method; 2) comparing this method to alternative approaches; and 3) showing its application to analyse the formation process of a site.

3.1 Relevance and reliability of the topological method

3.1.1 Consensus analysis supports the topological method

A first way to assess the relevance of a method is to compare its results with the intuitive reasoning of experts. Comparing the rankings of the eight theoretical graphs by archaeolo-

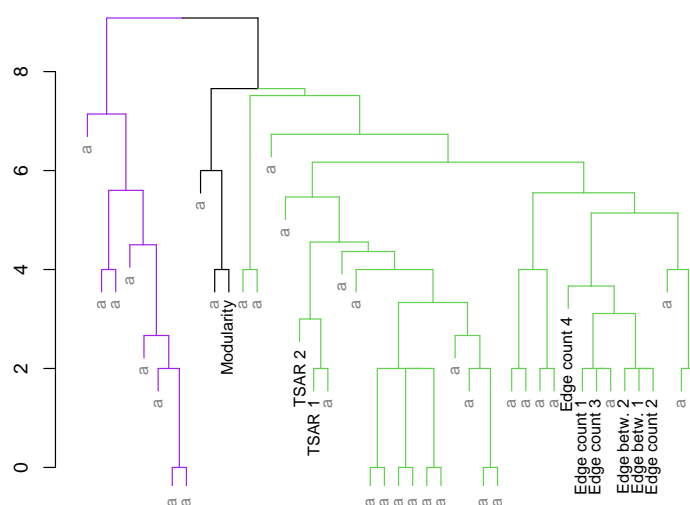


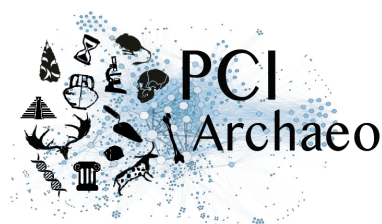
Figure 9. Clustering (UPGMA) of the distances (optimal matching) between the rankings of eight theoretical examples by archaeologists (labelled “a”) and by the four methods (TSAR admixture, TSAR admixture based on edge betweenness centrality, edge count, and modularity). The multiple results generated by some of the methods are numbered.

gists and the ranking generated by the topological admixture and the three other methods¹² showed (Figure 9): 1) an absence of consensus between archaeologists in ranking the patterns; 2) that the results can be grouped into three different clusters, with those from modularity in a specific cluster, and those from the other methods in a main cluster (in green in the dendrogram); and 3) within this main cluster, the results from the topological admixture are more similar to those from the survey than the results from the alternative methods, which are grouped in a specific subcluster. This supports the relevance of the topological admixture.

In conclusion, the discrepancy of the archaeologists' answers justifies the definition of an explicit method such as the TSAR method¹³, whose specificity is reflected in the results of this comparison.

¹²All analysis were made using the *archeofrag* R package (Plutniak, 2021), which is based on the *igraph* package (Csárdi and Nepusz, 2006). See supplementary material, section 4.1.2 “Comparison with archaeologists' rankings”.

¹³A similar discrepancy was reported in Fish, 1978, after comparing the classifications of 90 pottery sherds by four analysts.



3.1.2 Benchmark validates topological cohesion

Simulated graphs were used to benchmark the cohesion measure and test whether it adequately reflects both the effects of the relative size of the two layers and the effects of fragment movement. Sets of graphs were generated using different pairs of values for the “balance” and “disturbance” parameters. The cohesion values and the admixture of the two layers were measured on the generated graphs.

As evidenced by the results, similar admixture values can correspond to different proportions in the size of the layers (balance) (Figure 10). This observation justifies the need to not only rely on an admixture measure and supports the relevance of the cohesion measure implemented in the TSAR method. As expected, measuring cohesion distinguishes effectively between the effects of different size proportions (balance) and the effects of fragment movement between two layers (disturbance).

3.2 Comparison of methods

3.2.1 Ranking of the theoretical examples

Developing a new method is relevant only if it differs from previous methods and improves the description and comprehension of the phenomenon under study. To demonstrate the relevance of the TSAR approach, it was first applied to theoretical examples (Figure 8) and compared to the results from three other methods¹⁴ using a Spearman correlation test.

Results show that the admixture methods (both the TSAR and the edge betweenness variants) succeeded in discriminating the ten different theoretical cases (Figure 11). There is an exception for examples 4 and 8 which received the same value from the TSAR method; however, an intuitive evaluation would have been unable to discriminate this pair of graphs, given the complex structure of case 8. This result can be considered as a positive output of the method. Although the two variants of the topological admixture have an average (and not significant) correlation (Table 5), they differ due to a pair of fragments located in two different layers (example 1). The TSAR method ranks this case in the middle of the series, whereas the edge betweenness variant assigns a null admixture value. A null value is counter-intuitive and especially unsatisfying because isolated pairs of connected fragments are very frequent in archaeological observations.

The only statistically significant correlation is between the edge count and the modularity methods¹⁵. However, these methods do not distinguish between several sets of cases, notably between examples 1, 9, and 10 that are intuitively different and led to different archaeological interpretations. Finally, modularity is irrelevant since it reports distinct clusters in only two cases (examples 7 and 8).

¹⁴See supplementary material section 4.1.1 “Application of the four methods”.

¹⁵To turn modularity, a measure of distinction, into a measure of mixture, inverse values were used.

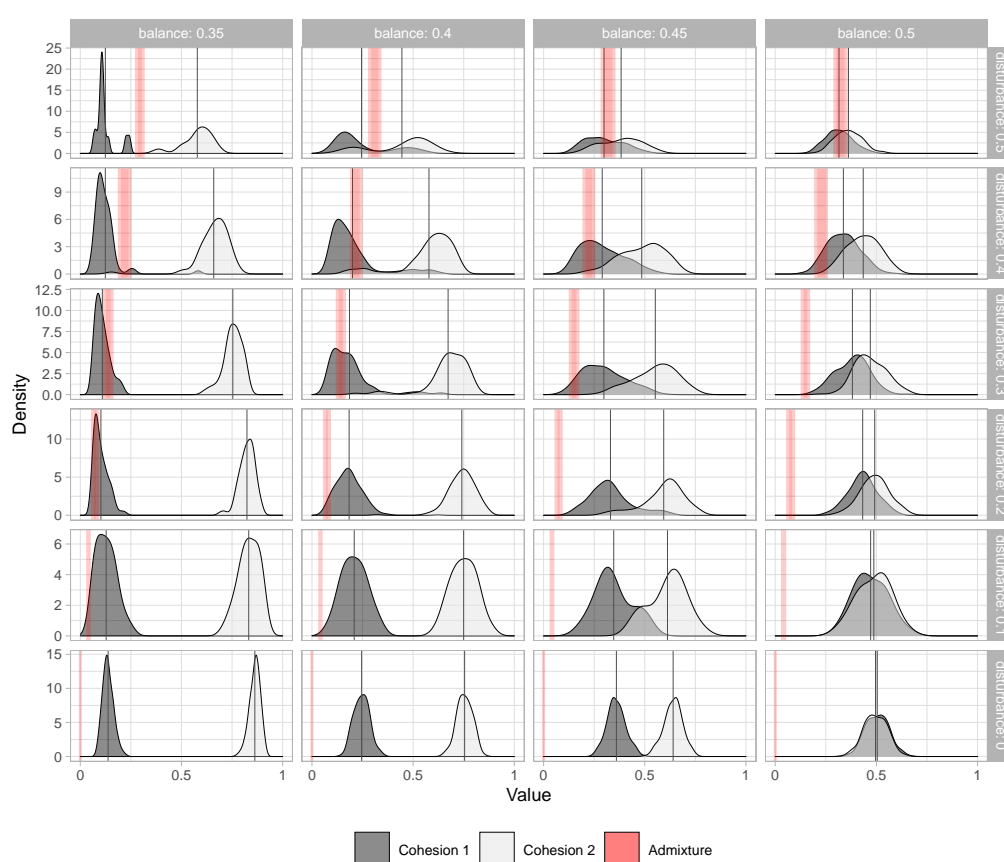


Figure 10. Density of the cohesion values for layers 1 and 2 of the simulated graphs (30 objects, 100 fragments) for different balance (columns) and disturbance (rows) parameters (6000 graphs). The black vertical lines give the mean cohesion values. Admixture is summarised by the red shades (representing the interquartile range of the values) and is also indicated by the positions of the cohesion values: the smaller the values, the more the layers are mixed. The reading is similar to the reading of Table 1. E.g., the upper left case corresponds to two layers with unequal cohesion values and high admixture.

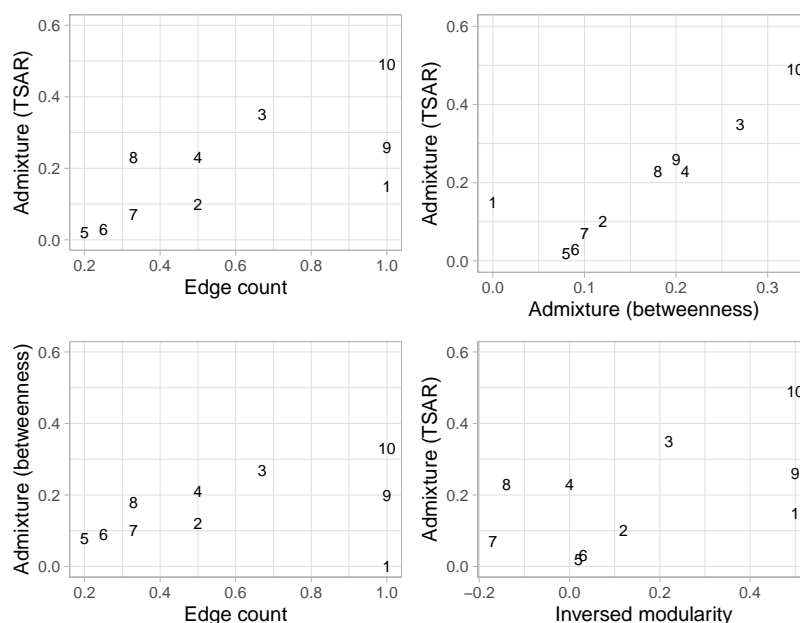


Figure 11. Comparison of the results of the four methods applied to the theoretical examples. The numbers refer to the labels of the graphs in Figure 8.

Table 5. Spearman correlation coefficients and [p-values] computed on the ranks of the ten theoretical examples sorted with the four methods of admixture measurement.

	Edge count	Admixture betw.	Modularity
Admixture betw.	0.30 [0.40]	–	–
Modularity	0.72 [0.02]	-0.13 [0.73]	–
Admixture TSAR	0.02 [0.96]	0.48 [0.16]	-0.09 [0.80]

3.2.2 Numerical differences between the methods

Simulated data were used to study in more detail the difference between the four methods. Absolute numerical differences between the values generated (between 0 and 1) is important because an archaeological interpretation relies on these differences (if a method gives an admixture of 0.45 and a different method gives 0.55, this difference of 0.1 might change the archaeological conclusion about the two considered layers).

Comparing the results from the edge count method and edge betweenness-based admixture to the results of the TSAR method reveals non-trivial differences, ranging from -0.27 to 0.21 (Figure 12). The edge betweenness-based admixture values are higher, especially when the disturbance (proportion of inter-layer relationships) increases. In contrast, edge count results are lower and stabilised for higher disturbance values.

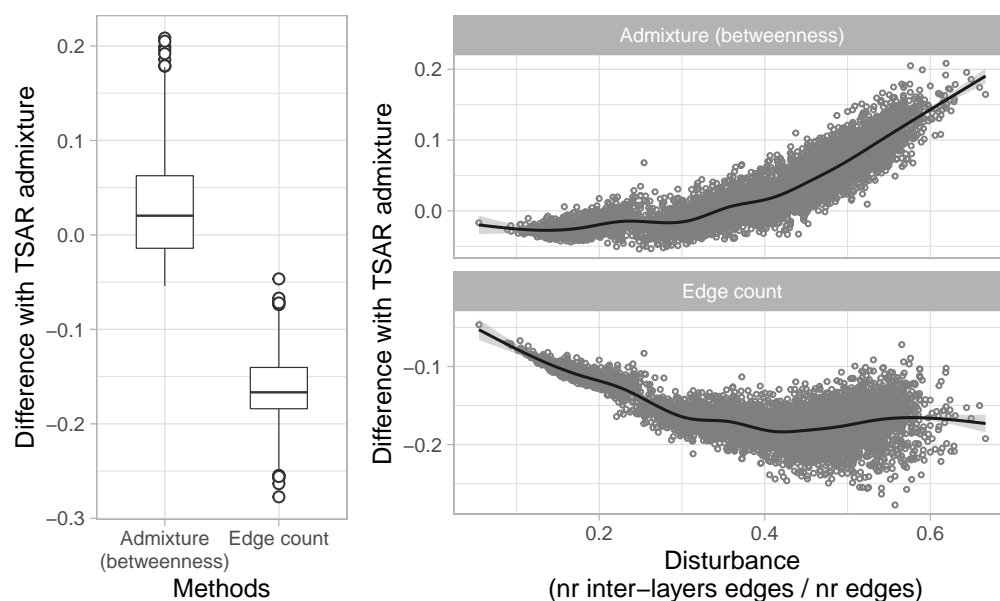


Figure 12. Differences between the values generated with the TSAR method and the two alternative methods (edge count and edge betweenness based admixture) for 1400 artificial graphs (4 replications each) by method (left) and as a function of the disturbance observed on these graphs (right).

In addition, a comparative analysis of the robustness of the edge count method and the TSAR admixture method was performed. First, a fragmentation graph with two layers was generated and its admixture was measured with both methods. These values were considered as the “true” admixture values of this graph. Second, the graph was altered by randomly removing a given proportion of edges (from 10% to 90%), simulating connection relationships not identified by the archaeologists. Admixture values were measured again and the difference with the “true” values were computed. Results of this procedure, repeated for 450 graphs, demonstrate that the TSAR method is much more robust and less sensitive to the lack of information¹⁶ (Figure 13).

These results demonstrate that: 1) some irrelevant results were generated by the alternative methods; and 2) the TSAR topological admixture is more sensitive to diverse archaeological situations, generating significantly different values; and 3) this method is more robust when information is missing, a common situation in archaeology.

¹⁶See supplementary material, section 4.3.3 “Robustness of the measurement methods”.

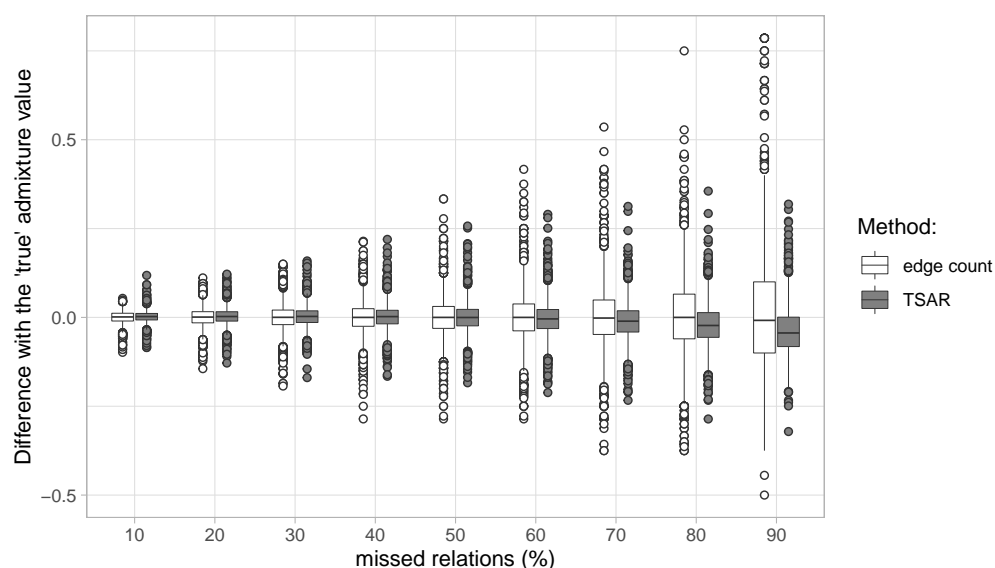


Figure 13. Differences between the “true” admixture value and the admixture values measured with two methods (edge count and edge betweenness based admixture) when simulating the non-observation of a given proportion of relationships (50 graphs generated for each proportion value).

3.3 Testing the formation process hypotheses at Liang Abu

3.3.1 Evaluating layers at Liang Abu

The four methods were applied to layers 0 (the actual surface), 1, and 2 of Liang Abu (Table 6). The TSAR method reports an admixture value equal to 0 for layers 0 and 1 (as expected since there is no connection relationship between them), and a low value equal to 0.01 for layers 1 and 2. Alternative methods report higher admixture values for layers 1 and 2. Comparing the relative range of values generated by each method shows that the TSAR method gives a lower difference in the robustness of the distinctions between layer 0 and 1, and 1 and 2, respectively. Note that modularity adequately suggests a distinction between both pairs of layers, with a distinction between layers 0 and 1 that is weaker than between layers 1 and 2. However, we must note that modularity is based on irrelevant assumptions in this archaeological context.

In addition, as demonstrated above from the simulated data, admixture alone can be ambiguous and must be complemented by an examination of the cohesion values. Layer 1 appears much more cohesive in the comparison of layers 0 and 1 than in the comparison of layers 1 and 2, because layer 1 has a significantly lower cohesion value in the former comparison. Referring to the interpretation grid (Table 1), we conclude that: 1) layers 0 and 1 corre-



spend to case 3 (significant difference in the cohesion value and low admixture), suggesting that only layer 1 is validated; and 2) layers 1 and 2 correspond to case 4 (minor difference in the cohesion value and low admixture), leading to a validation of both layers.

Comparing the results from different pairs of layers from the same site (or from different sites) can make these numbers meaningful and helpful for archaeological interpretation. However, this empirically-based comparison is limited by the quantity of data available from a given site(s), and by our ignorance of the past states of the site, making it impossible to determine the deposition process. Using simulated data will overcome these limitations and refine the meaning of these numerical results.

Table 6. For each pair of pottery layers at Liang Abu (0 and 1, 1 and 2), the table reports the number of sherds, the maximal number of single objects the sherds come from, the TSAR cohesion and admixture values, and the values of the alternatives methods.

	Layers 0 & 1	Layers 1 & 2
Objects	11	28
Fragments	29	72
Cohesion layer 0	0.09	–
Cohesion layer 1	0.91	0.40
Cohesion layer 2	–	0.59
Admixture (TSAR)	0.00	0.01
Admixture (betweenness)	0.00	0.02
Edge count	0.00	0.06
Modularity	0.30	0.41

3.3.2 Comparing with simulated data and testing the formation hypotheses

Simulation is used for two purposes: first, to study the relationship between the properties of an empirical fragmentation graph and similar simulated graphs; second, to test different scenarios of the formation processes which led to the observed archaeological situation. Given two spatial units with a non-null admixture value, two hypotheses about the distinction between these units were considered: H1, the archaeological material studied comes from a single deposition episode, within which archaeologists distinguished two subsets; H2, the material was deposited during two deposition episodes, that archaeologists could not distinguish due to subsequent perturbations, admixture, and sampling. Each hypothesis has different consequences for archaeological interpretation. With H1, the distinction between the two layers must be abandoned and therefore the idea of their admixture. With H2, the distinction between the two layers can be conserved and their admixture is explained by the movements of fragments.

The fragmentation process was run with two different initial conditions, assuming or not a boundary between the locations of the two sets of objects being fragmented (Figure 7).The

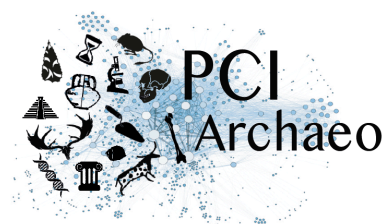


Table 7. Interpretation table for the comparison between the empirical values and the simulated values for H1 and H2, as a function of: 1) the relationship between the two distributions of simulated values; and 2) the inclusion or exclusion of the empirical values in these distributions.

		Simulated distributions for H1 and H2	
		similar	different
observed value	included	expected situation	confirmation of H1 or H2
	excluded	strong anomaly	anomaly

properties of the graphs generated for each hypothesis were then compared to the empirical values to determine the most likely hypothesis. Two aspects are considered to interpret the numerical results:

- whether the empirical values of interest were excluded or included in the H1 or H2 distributions (or both), by visual estimation, using the interquartile range of the simulated value;
- whether the simulated values for H1 and H2 were significantly different or not, determined by visual estimation and by testing the difference in the median values with a Wilcoxon signed-rank test.

Considering a parameter (e.g., admixture), if the values for H1 and H2 significantly differ and if the empirical value is similar to one of those values, then it supports the corresponding formation hypothesis (Table 7). It must be stressed that, based on these numerical instruments and results, the final interpretation is qualitative and left to the archaeologist who uses their general knowledge of the archaeological context.

This approach was applied to layers 1 and 2 from Liang Abu. The simulator was run twice, for the initial condition H1 and H2. Apart from the number of initial layers, the other parameters are inferred from the properties of the fragmentation graph corresponding to layers 1 and 2 from Liang Abu¹⁷. This means that, for example, one assumes that the balance observed empirically reflects the balance at the “initial” state of the two layers in the past. A different approach would have been to set the balance of the simulator with the proportion of pottery material recovered from the two layers (defined from the number or weight of the pieces). The several parameters of the TSAR simulator enable us to test multiple hypotheses, but we keep it simple for this general presentation.

Using the interpretation grid in Table 7, we observe that for the edge count, edge weight sum, and the balance parameters, the empirical and simulated results are rather similar and do not give clear conclusions. A strong anomaly is revealed for the disturbance, the empirical

¹⁷See supplementary material, section 4.4.2 “Testing formation hypotheses for Liang Abu layers 1 and 2”.

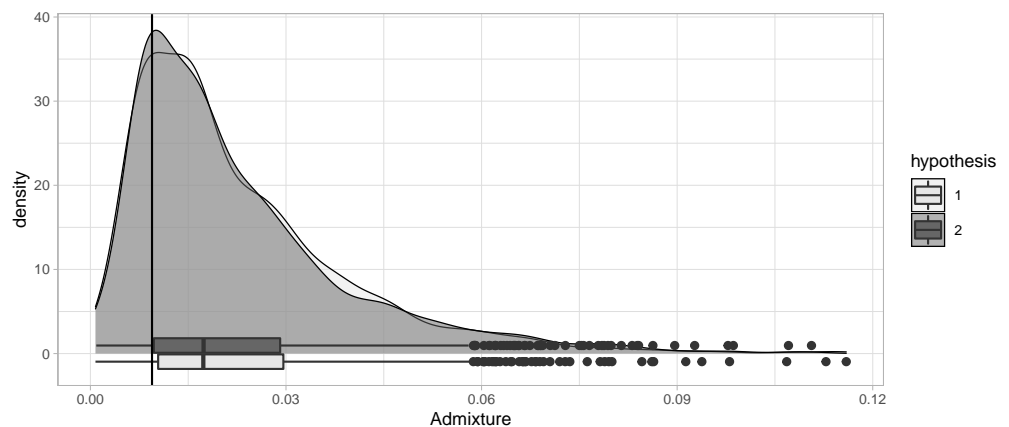


Figure 14. Admixture observed on graphs generated using Liang Abu layers 1 and 2 parameters for two formation process hypotheses (1000 replications each), compared with the admixture observed at Liang Abu (vertical bar).

value is interestingly much lower than the values from both hypotheses, which do not significantly differ. Conversely, the empirical admixture value is the same as the expected value from the simulation (Figure 14). Concerning cohesion, the difference in the median of values of H1 and H2 is supported by the Wilcoxon test. In addition, the empirical cohesion values show a clear agreement with H2, whereas they are outside the interquartile ranges of the results simulated for H1 (Figure 15). This suggests favouring the scenario where layers 1 and 2 were initially two independent layers. Therefore, the analysis of pottery fragmentation and refitting validates the distinction between these layers at Liang Abu.

This case study of a small data set illustrates how the TSAR method and simulation can enhance fragmentation analysis in archaeological contexts. In future work, this approach should be applied and tested on larger datasets using all the potential of the simulator with different settings to test more sophisticated formation hypotheses.

4 Conclusion

This paper presented a renewed framework for “refitting” analysis in archaeology. Using graph theory to model the topological relationships between fragments that were formerly parts of the same initial object can define new measurements to evaluate the reliability of archaeological spatial units (such as stratigraphic layers). This approach requires a more time-consuming recording method but generates more accurate and substantiated results. In addition, the development of this formal framework for refitting and stratigraphic analysis paves the way for conceptual clarifications. This has led, in particular, to redefining “refitting” as *connection*, to approximate a definition of a *layer* and other *archaeological spatial units*, and

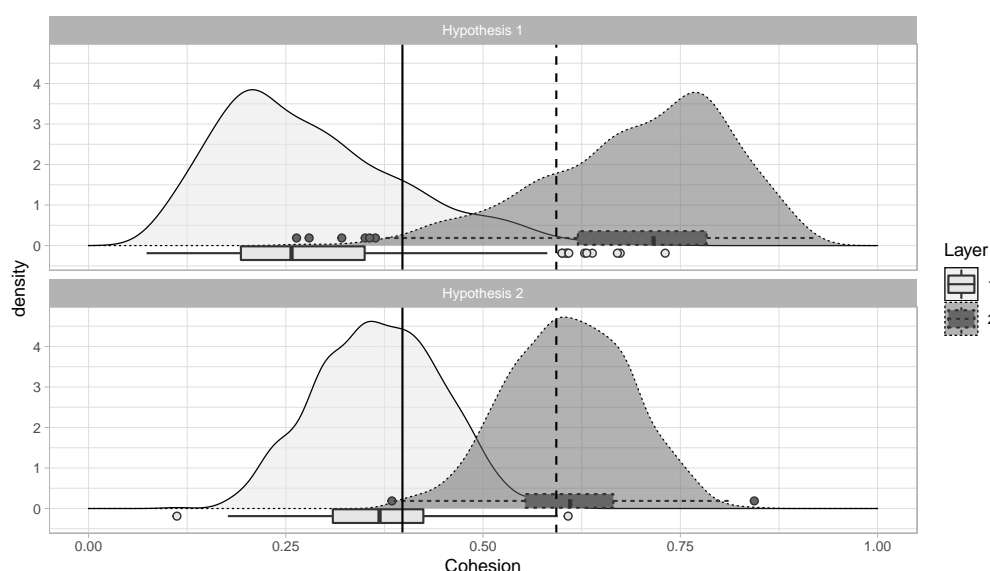


Figure 15. Cohesion values observed on graphs generated using Liang Abu layers 1 and 2 parameters for two formation process hypotheses (1000 replications each), compared with the cohesion values observed at Liang Abu (vertical bar).

to model the *fragmentation process*.

However, more research has to be done in this direction, addressing fundamental archaeological concepts using concepts and tools from the fields of formal and applied ontology. This is an indirect, albeit essential, result of this early research. Further methodological developments will concern: 1) applying the method to larger data sets; 2) weighting cohesion and admixture with morphometric values (e.g., sherd size for pottery) and spatial distances¹⁸; and 3) using the topological properties of connected fragment networks as a proxy to detect technological features or human behaviour (e.g., intentional breaking). In compliance with the principles of reproducible and reusable research, and in particular, with the idea that scientific scholarship should be embedded as software (Donoho et al., 2008), the supplementary material of this paper takes the form of an “executable paper” (Leisch et al., 2011), and the TSAR method as an *Archeofrag* R package¹⁹. This package is complemented by a *Shiny* application²⁰, encouraging its demonstration, use, and dissemination.

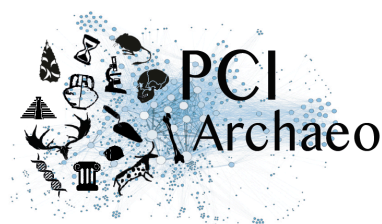
Data accessibility

Data are available online: DOI: 10.5281/zenodo.4271900.

¹⁸This feature is implemented in the *archeofrag* package from v0.7.

¹⁹Available on CRAN at <https://cran.r-project.org/package=archeofrag>.

²⁰<https://analytics.huma-num.fr/Sebastien.Plutniak/archeofrag>.



Supplementary material

Script and codes are available online: DOI: [10.17605/osf.io/9c8ge](https://doi.org/10.17605/osf.io/9c8ge).

Acknowledgements

I thank all the participants to the survey, Robert Bischoff, Matthew Peeples and the anonymous reviewer who evaluated this paper for the *PCI Archaeology*, Juan A. Barceló, Guillaume Cabanac, Joséphine Caro, Julian Laabs, Claire Manen, Luce Prignano, Oliver Nakoinz, and Émilien Schultz for their valuable comments at different stages in the development of this research. I also acknowledge François-Xavier Ricaut and the MAFBO (<https://kaltim.hypotheses.org>) and OceoAdapto (ANR-14-CE31-0013-01 grant) projects, which made the excavation of Liang Abu possible. The finalisation of this paper was made during a stay at the *Institut für Ur- und Frühgeschichte* of Kiel, thanks to the support of a “Short-term research grant” from the *Deutscher Akademischer Austausch Dienst* (DAAD). Version 3 of this preprint has been peer-reviewed and recommended by Peer Community In Archaeology (<https://doi.org/10.24072/pci.archaeo.100010>).

5 Conflict of interest disclosure

The author of this preprint declare that he has no financial conflict of interest with the content of this article. Sébastien Plutniak is recommender for PCI Archaeology.

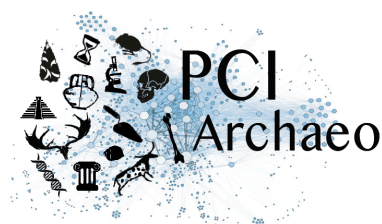
References

- Barthès P (1994). L'apport du remontage des vases dans l'analyse et l'interprétation d'une stratigraphie d'habitat néolithique. L'exemple de l'abri Font-Juvénal à Conques-sur-Orbiel (Aude). In: *Terre cuite et société. La céramique, document technique, économique, culturel. XIVe Rencontres internationales d'archéologie et d'histoire d'Antibes* (Antibes, Oct. 21–23, 1993). Juan-les-Pins: éditions APDCA, pp. 199–215. isbn: 9782904110177.
- Bollong CA (1994). Analysis of Site Stratigraphy and Formation Processes using Patterns of Pottery Sherd Dispersion. *Journal of Field Archaeology* 21, 15–28. doi: [10.1179/009346994791549254](https://doi.org/10.1179/009346994791549254).
- Bordes JG (2000). La séquence aurignacienne de Caminade revisitée: l'apport des raccords d'intérêt stratigraphique. *Paléo. Revue d'archéologie préhistorique* 12, 387–407. doi: [10.3406/pal.2000.1611](https://doi.org/10.3406/pal.2000.1611).
- Brantingham PJ, TA Surovell, and NM Waguespack (2007). Modeling Post-depositional Mixing of Archaeological Deposits. *Journal of Anthropological Archaeology* 26, 517–540. issn: 0278-4165. doi: [10.1016/j.jaa.2007.08.003](https://doi.org/10.1016/j.jaa.2007.08.003).



- Cahen D and J Moeyersons (1977). Subsurface Movements of Stone Artefacts and their Implications for the Prehistory of Central Africa. *Nature* 266, 812–815. doi: 10.1038/266812a0.
- Caron F, F d'Errico, P Del Moral, F Santos, and J Zilhão (2011). The Reality of Neandertal Symbolic Behavior at the Grotte du Renne, Arcy-sur-Cure, France. *Plos One* 6, 1–11. doi: 10.1371/journal.pone.0021545.
- Clauset A, MEJ Newman, and C Moore (2004). Finding Community Structure in Very Large Networks. *Physical Review E* 70 (6), 066111. doi: 10.1103/PhysRevE.70.066111.
- Csárdi G and T Nepusz (2006). The igraph Software Package for Complex Network Research. *InterJournal* 1695, 1–9.
- Cziesla E (1990). On Refitting of Stone Artefacts. International Symposium on Refitting Stone Artefacts. In: *The Big Puzzle* (Monrepos, Sept. 6–9, 1987). Ed. by Cziesla E, Eickhoff S, Arts N, and Winter D. Studies in Modern Archaeology 1. Bonn: Holos, pp. 9–44.
- Cziesla E, S Eickhoff, N Arts, and D Winter, eds. (1990). *The Big Puzzle. International Symposium on Refitting Stone Artefacts* (Monrepos, Sept. 6–9, 1987). Studies in Modern Archaeology 1. Bonn: Holos.
- Donoho D, A Maleki, M Shahram, I Rahman, and V Stodden (2008). *15 Years of Reproducible Research in Computational Harmonic Analysis*.
- Fish PR (1978). Consistency in Archaeological Measurement and Classification: A Pilot Study. *American Antiquity* 43, 86–89. doi: 10.2307/279635.
- Fortunato S and M Barthélemy (2007). Resolution Limit in Community Detection. *Proceedings of the National Academy of Sciences* 104, 36–41. issn: 0027-8424. doi: 10.1073/pnas.0605965104.
- Girvan M and MEJ Newman (2002). Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences* 99, 7821–7826. doi: 10.1073/pnas.122653799.
- Gnaden D and S Holdaway (2000). Understanding Observer Variation When Recording Stone Artifacts. *American Antiquity* 65, 739–747. doi: 10.2307/2694425.
- Gray L (2003). A Mathematician looks at Wolfram's New Kind of Science. *Notices of the American Mathematical Society* 50, 200–211.
- Harris EC (1979). *Principles of Archaeological Stratigraphy*. London: Academic Press. xiii–136 p.
- Hofman JL (1986). Vertical Movement of Artifacts in Alluvial and Stratified Deposits. *Current Anthropology* 27, 163–171. doi: 10.1086/203414.
- (1992). Piecing Together the Past: an Introduction to Refitting. In: *Piecing Together the Past. Applications of Refitting Studies in Archaeology*. Ed. by Hofman JL and Enloe JG. British Archaeological Reports International Series 578. Oxford: Tempus Reparatum, pp. 1–20. isbn: 978-0860547334.
- Hofman JL and JG Enloe, eds. (1992). *Piecing Together the Past. Applications of Refitting Studies in Archaeology*. British Archaeological Reports International Series 578. Oxford: Tempus Reparatum. 321 pp. isbn: 978-0860547334.

- Johnson DL (1989). Subsurface Stone Lines, Stone Zones, Artifact-Manuport Layers, and Bioman-
tles Produced by Bioturbation via Pocket Gophers (Thomomys Bottae). *American Antiquity*
54, 370–389. doi: 10.2307/281712.
- Leisch F, M Eugster, and T Hothorn (2011). Executable Papers for the R Community: The R2
Platform for Reproducible Research. *Procedia Computer Science* 4, 618–626. doi: 10.1016/
j.procs.2011.04.065.
- López-Ortega E, XP Rodríguez, and M Vaquero (2011). Lithic Refitting and Movement Connec-
tions: the NW area of level TD10-1 at the Gran Dolina site (Sierra de Atapuerca, Burgos,
Spain). *Journal of Archaeological Science* 38, 3112–3121. doi: 10.1016/j.jas.2011.07.011.
- Lyman RL and MJ O'Brien (1999). Americanist Stratigraphic Excavation and the Measurement
of Culture Change. *Journal of Archaeological Method and Theory* 6, 55–108. doi: 10.1023/A:
1022976726447.
- May K (2020). The Matrix: Connecting Time and Space in Archaeological Stratigraphic Records
and Archives. *Internet Archaeology* 55. doi: 10.11141/ia.55.8.
- Michel R (2002). *Saint-Blaise/Bains des Dames, 3. Typologie et chronologie de la céramique néolithique.*
Céramostratigraphie d'un habitat lacustre. 2 vols. Archéologie neuchâteloise 27. Neuchâtel:
Service et musée cantonal d'archéologie. 222 pp. isbn: 978-2940347230.
- Moberg CA (1971). Mathematics in the Archaeological and Historical Sciences. In: *Archaeo-
logical Context and Mathematical Methods. Proceedings of the Anglo-Romanian Conference*
(Mamaia, Sept. 16–24, 1970). Ed. by Hodson FR, Kendall DG, and Tăutu P. Edinburgh: Ed-
inburgh University Press, pp. 551–565. isbn: 0-85224-213-1.
- Morin E, T Tsanova, N Sirakov, W Rendu, JB Mallye, and F Lévêque (2005). Bone Refits in Strat-
ified Deposits: Testing the Chronological Grain at Saint-Césaire. *Journal of Archaeological
Science* 32, 1083–1098. doi: 10.1016/j.jas.2005.02.009.
- Myers OH (1958). 'Abka Re-excavated. *Kush* 6, 131–143.
- Newman MEJ (2006). Modularity and Community Structure in Networks. *Proceedings of the
National Academy of Sciences* 103, 8577–8582. doi: 10.1073/pnas.0601602103.
- O'Brien MJ and RL Lyman (2002). The Epistemological Nature of Archaeological Units. *Anthro-
pological Theory* 2, 37–56. doi: 10.1177/1463499602002001287.
- Ore Ø (1962). *Theory of Graphs*. Providence (R.I.): American Mathematical Society. 270 pp.
- Plutniak S (2021). *Archeofrag: an R package for Refitting and Spatial Analysis in Archaeology*. R
package version 0.7. doi: 10.5281/zenodo.4271900.
- Plutniak S, A Araujo, S Puaud, JG Ferrié, AA Oktaviana, B Sugiyanto, JM Chazine, and FX Ricaut
(2016). Borneo as a Half Empty Pot: Pottery Assemblage from Liang Abu, East Kalimantan.
Quaternary International 416: *Southeast Asia: Human Evolution, Dispersals and Adaptation*,
228–242. issn: 1040-6182. doi: 10.1016/j.quaint.2015.11.080.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for
Statistical Computing. Vienna, Austria.



- Rowlett RM and MC Robbins (1982). Estimating Original Assemblage Content to Adjust for Post-depositional Vertical Artifact Movement. *World Archaeology* 14, 73–83. doi: 10.1080/00438243.1982.9979850.
- Schaller-Åhrberg E (1990). Refittings as a Method to Separate Mixed Sites: a Test with Unexpected Results. International Symposium on Refitting Stone Artefacts. In: *The Big Puzzle* (Monrepos, Sept. 6–9, 1987). Ed. by Czesla E, Eickhoff S, Arts N, and Winter D. Studies in Modern Archaeology 1. Bonn: Holos, pp. 611–622.
- Schurmans U (2007). Refitting in the Old and New Worlds. In: *Fitting Rocks. Lithic Refitting Examined*. Ed. by Schurmans U and De Bie M. British Archaeological Reports International series 1596. Oxford: Archaeopress, pp. 7–23. isbn: 978-1-4073-0012-2.
- Schurmans U and M De Bie, eds. (2007). *Fitting Rocks. Lithic Refitting Examined*. British Archaeological Reports International series 1596. Oxford: Archaeopress. isbn: 978-1-4073-0012-2.
- Sisk ML and JJ Shea (2008). Intrasite Spatial Variation of the Omo Kibish Middle Stone Age Assemblages: Artifact Refitting and Distribution Patterns. *Journal of Human Evolution* 55, 486–500. doi: 10.1016/j.jhevol.2008.05.016.
- Surovell TA, NM Waguespack, JH Mayer, M Kornfeld, and GC Frison (2005). Shallow Site Archaeology: Artifact Dispersal, Stratigraphy, and Radiocarbon Dating at the Barger Gulch Locality B Folsom Site, Middle Park, Colorado. *Geoarchaeology* 20, 627–649. doi: 10.1002/gea.20071.
- Villa P (1982). Conjoinable Pieces and Site Formation Processes. *American Antiquity* 47, 276–290. doi: 10.2307/279901.
- Villa P and J Courtin (1983). The Interpretation of Stratified Sites: a View from Underground. *Journal of Archaeological Science* 10, 267–281. doi: 10.1016/0305-4403(83)90011-0.
- Wasserman S and K Faust (1994). *Social Network Analysis. Methods and Applications*. Cambridge: Cambridge University Press. xxxi–825 p. isbn: 0-521-38269-6.
- Whittaker JC, D Caulkins, and KA Kamp (1998). Evaluating Consistency in Typology and Classification. *Journal of Archaeological Method and Theory* 5, 129–164. doi: 10.1007/BF02427967.
- Wood WR and DL Johnson (1982). A Survey of Disturbance Processes in Archaeological Site Formation. In: *Advances in Archaeological Method and Theory. Selections for Students from Volumes 1–4*. Ed. by Schiffer MB, pp. 539–605. isbn: 978-0-12-624180-8. doi: 10.1016/B978-0-12-624180-8.50016-0.
- Ziesaire P (1990). Refitting Flaking Tools from an Early Mesolithic site at Altwies-Haed, Grand Duchy of Luxemburg. International Symposium on Refitting Stone Artefacts. In: *The Big Puzzle* (Monrepos, Sept. 6–9, 1987). Ed. by Czesla E, Eickhoff S, Arts N, and Winter D. Studies in Modern Archaeology 1. Bonn: Holos, pp. 253–261.