The strength of parthood ties. Evaluating archaeological layers using graph theory to model objects refitting

Sébastien Plutniak

Traces laboratory, Toulouse, France

Contents

Abstract

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

Keywords: refitting, graph analysis, network analysis, stratigraphy, post-depositional process, taphonomy, software

1 Introduction

Modern studies of archaeological stratigraphy involve numerous specialists, each of them shedding specific light on the relevant distinctions to discretise 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 identify 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 which were parts of the same original object, has been long-used as a method to assess the integrity of archaeological discrete spatial units[1]. Inter-layers displacement of remains are used as clues of post-depositional disturbances (Myers 1958, Villa 1982, Hofman 1986, Ziesaire 1990, Barthès 1994, Bordes 2000, Morin et al. 2005), intra-sites refitting are interpreted as pieces of evidences of site re-use (Cahen and Moeyersons 1977), and inter-sites admixture are considered as shreds of evidences of people movement (Schaller-Åhrberg 1990).

In the last three decades, refitting analysis has become an extensively studied field, addressed through several conferences and collective publications. This scholarship has generated numerous methodological improvements, which cannot be summarised here (Cziesla et al. 1990, Hofman and Enloe 1992, Schurmans and De Bie 2007; for an history, see Schurmans 2007). Post-depositional processes such as trampling (Villa and Courtin 1983) and movements have been long-recognised among possible disturbance processes (Wood and Johnson 1982, Johnson 1989). Accordingly, about thirty years ago, Hofman urged archaeologists to not simply use refitting to acknowledge disturbances occurred, but rather to document "the degree of such movements", namely to quantify them (Hofman 1992, p. 4). In this regard, many methods were 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 seven types, also about lithic artefacts (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 to quantify the number of relations alone can be misleading and should be complemented by quantification of the "structure" of these relations, i.e. the *topology* of the refitting relations between pieces. Note that considering

---

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

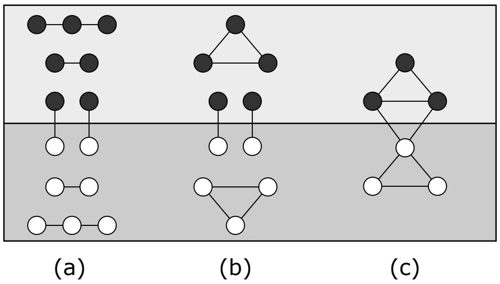

Figure 1: The need to consider topology: three examples of two layers with internal refitting (n=6) and inter-layers refitting (n=2). Although the numbers of relations are equal in all examples, their archaeological interpretation would be very different. The distinction between the two layers would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

the topological aspect in this context is not a completely new idea (Cziesla 1990, Michel 2002). However, previous attempts were limited to manual graphic procedures, based on the count of (different types of) relations, and were not generalised and implemented into actual data processing procedures. There are indeed good reasons to integrate the topology of refitting relations, as illustrated by the theoretical examples in Figure 1: although the count of refits and the proportion of internal versus external relations are equal, conclusions about the validity of the layers and the initial location of the remains would be different. To overcome these ambiguities, graph theory will be use to model sets of refitting relations and study their topology. The following of this paper will aim to 1) define and implement the Topological Study of Archaeological Refitting (TSAR), a method to quantify the "strength" of archaeological units (layers), based on the topology of parthood relations, i.e. relations between parts and wholes; 2) compare TSAR with previous and alternative procedures; 3) present results using simulated data and an application to field archaeological data.

2 Material and methods

2.1 Preliminary definitions

Before going into more details about the design and implementation of the TSAR method, some preliminary definitions are required.

2.1.1 Layers and spatial units

The problem addressed by the topological method initially relates to stratigraphical analysis (Harris 1979, Lyman and O'Brien 1999, May 2020). Archaeological observation proceeds 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 an example of archaeological discrete spatial unit, characterised by its special emphasis on vertical orientation and related by binary relationships (such as *being above*, *being below*, and *being inside*). In what follows, 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 other sort of archaeological spatial units as long as they contain fragmented objects (e.g. between the inside and outside of holes, buildings, specialised area, etc.).

2.1.2 Connection and similarity relationships

Consider a fragmented material object. All fragments share the same parthood relation with the initial object, of which they are parts. However, in this archaeological context, two types of relationships between fragments are distinguished, namely connection and similarity[2]. Connection relationship is taken as primitive relation in the TSAR approach. This term is used as a shorthand referring to the connection relationship which existed in the past between the (future) two fragments, before the object they composed was broken (Figure 2). Archaeologists deduce this relation from the symmetry between regions of the surface of the fragments, which can be physically adjusted (the fragments "refit"). The second type of relation, "similarity", concerns fragments considered as sharing enough common features (motif, clay, inclusions, etc.) to state they are parts of the same initial object. Connection and similarity relations have very different epistemic properties. Connection relationships rely on morphological features determined by solid-state physics, whereas the similarity relies on various analytical procedures used by archaeologists to determine similarity: perception of colour and texture, stylistic analysis, chemical determination, etc. Accordingly, similarity relationships are a matter of judgment. It is noteworthy that the related features have different meanings depending on the context they are used: for example, the similar chemical composition can prove two obsidian pieces come from the same volcano, but cannot prove undoubtedly that two pottery sherds come from the same pot.

2.1.3 Topological properties

Contiguity Determining contacts between fragments raises a practical difficulty which is also related to a classic issue regarding contiguity. In practice, the contact between two fragments on a too small surface makes ambiguous the determination of their connection. Conceptually, two types of contiguity are distinguished in spatial analysis, namely "Queen" and "Rook" contiguities[3]. Queen contiguity refers to contiguity determined by the sharing of at least a point (a line in 3D), whereas Rook contiguity is determined by the sharing of a line (a surface in 3D). The choice of the type of contiguity has important effects on subsequent computations: consider a chessboard, a cell has 9 neighbours according to Queen contiguity and only 4 according to Rook contiguity.

---

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

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

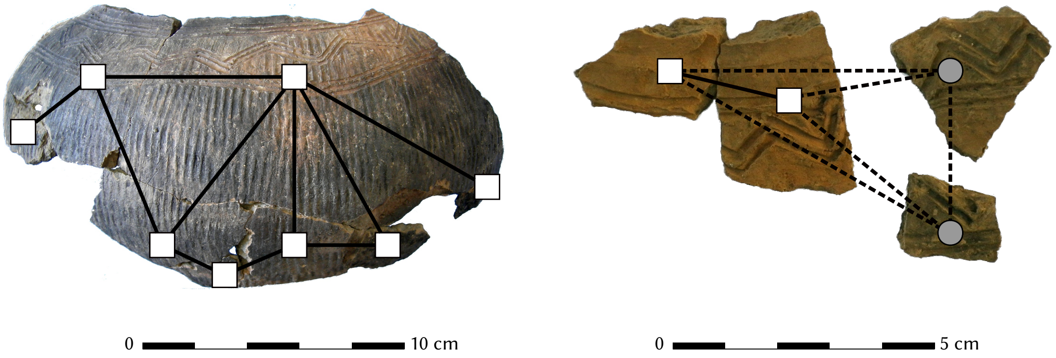

Figure 2: Pottery sherds from Liang Abu related by connection relationships (solid line between white squares) and similarity relationships (dashed lines between grey circles).

In the context of the analysis of archaeological fragments, given that connection relations are determined from the examination of surfaces, Rook contiguity must be favoured due to this conceptual correspondence, but also for practical and reliability reasons (at the macro, daily-life, scale of analysis used in archaeological refitting analysis, a "point-like" or "line-like" contact is not enough to determine a former connection between two fragments).

Graph planarity The concepts and tools of graph theory can be used to model the connection relationships of archaeological fragments. This implies to define the type of graph required, e.g. using undirected edges (since connection relationship is symmetric). In the context of the relation 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 led to identifying that some specific sets of connecting fragments have to be represented by non-planar graphs. This would be rather rare in the case of pottery (see Figure 3) 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 relations identified, will all determine the likelihood to require or not non-planar graphs to model the connection relationships between fragments (e.g. palaeolithic stone artefacts, sparse neolithic simple-shaped potteries, numerous medieval urban vessel with complex shapes).

---

[4] Note 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 making of the object by humans.

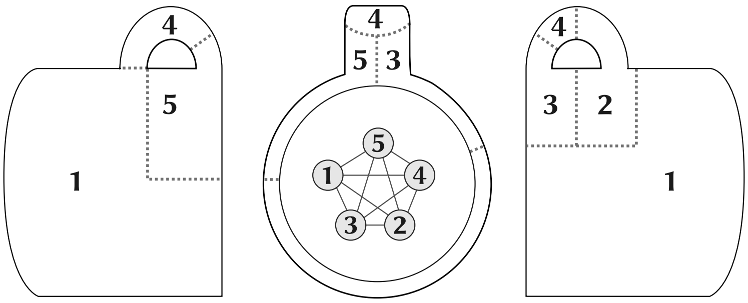

Figure 3: Non-planar fragmentation graph. Virtual example of a pot with a handler fragmented into five pieces. The corresponding non-planar graph is represented in the middle of the figure.

2.2 Measures

2.2.1 Principles of the topological method

Taking topology into account is the main principle of the TSAR method. However, related conceptual choices must now be clarified.

Type of relations Two types of relationships between fragments were distinguished, connection and similarity. However, this first version of the TSAR method only concerns connection. Due to their different epistemic and conceptual properties to integrate them into the same model makes the mathematics much more complex. This is due, in particular, to two main reasons: similarity relationships imply to account for degrees of certainty, and similarity is a transitive relation (A is similar to B, B to C, so C to A) whereas connection is not.

Consideration of the samples 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. Assuming that the reliability of the determination of a layer is strengthened as it contains more material is acceptable archaeological reasoning. However, the TSAR method is only about remains concerned by connection relationships. Fragments without connection, "singletons", are excluded. Therefore, the interpretation of the result in consideration to the total quantity of remains is l~~et~~ to the archaeologist. Methodological coherence justifies this choice. It is grounded on the principle that an object or a fragment had a unique location before the fragmentation and dispersion processes happened. Fragmentation analysis is about observable archaeological evidence of those processes. Singletons are weak proofs of their initial location since they are only related to their layers (by a relation of inclusion). To the contrary, more fragments sharing a similar location are connected, 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 fragments and connection relationships. This approach enables to distinguish 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

	Difference between the cohesion values	
Admixture	−	+
+	1. one layer	2. movements from one layer
−	3. movements between two layers	4. movements between two layers

Table 1: Possible interpretations about the formation process of two layers as a function of the magnitude of their admixture and of the difference between their cohesion values. "Movements" refers to transport of fragments between layers. See also illustrations in Figure 4.

relationships are observed between two layers containing 50 fragments each.

Layers admixture and reliability The final aim of the TSAR method is to evaluate the reliability of a distinction between two layers and to interpret the site formation process. Achieving this implies to evaluate and to compare the degree in which the layers are self-adherent to themselves, their cohesiveness. The opposite of layers' cohesion is their admixture. The concept of admixture supposes, first, to distinguish two containers and their respective content (layers and fragments) and, second, to determine that a part of the content is not within its expected container (fragments which moved to a different layer). Considering two layers, the TSAR method proceeds by defining a virtual third layer including the fragments and connection relationships at the intersection of the two layers. The admixture corresponds therefore to the part of the global cohesion which is not specific to a layer.

Cohesion and admixture are the two fundamental concepts to interpret the reliability of a distinction between two layers and the post-depositional processes which might have disturbed them. Keeping these concepts separated in the analysis enables to distinguish the different factors at work and propose different interpretations. Observing two layers with similar cohesion values and low admixture suggests to validate the distinction between these layers and to state that some movement of fragments happened (case 3 of Table 1 and Figure 4). Two layers with very different cohesion values but low admixture would lead to a similar conclusion (case 4). To the contrary, the distinction between the two layers would be considered as irrelevant when two layers with similar cohesion values present a high admixture (case 1). Finally, in the case of a high admixture between two layers with very different cohesion values, the determination of the layer with the highest cohesion would be validated and movements of fragments from this layer would be suggested, while the recognition of the other would remain uncertain (case 2).

⁂

The TSAR method includes three steps: 1) weighting the connection relationships, 2) measuring the cohesion of each layer and the admixture of the two layers, 3) interpreting the reliability of the distinction between the layer and the possible disturbance based on the cohesion and admixture values.

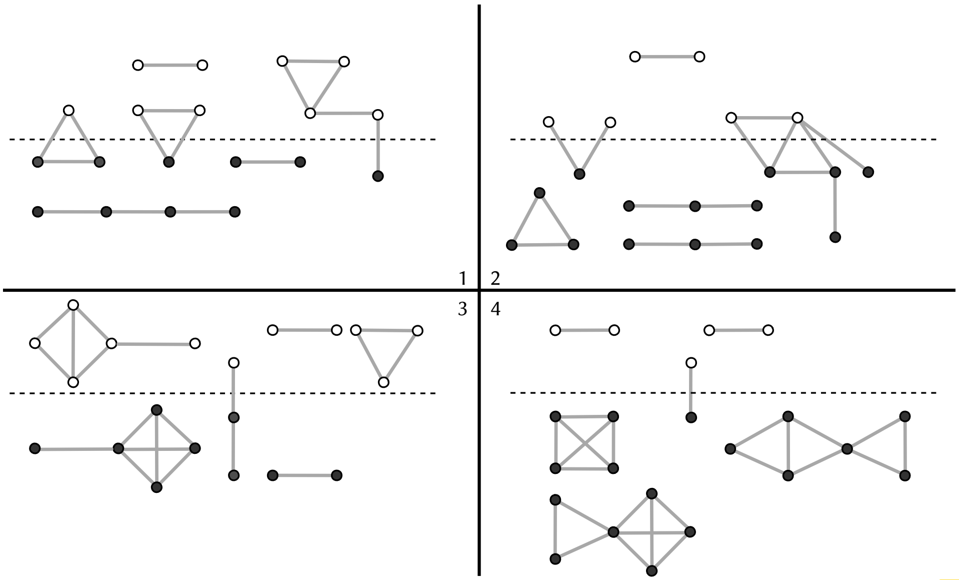

Figure 4: Illustrations for the four possible interpretations presented in Table 1, generated using the TSAR simulator.

2.2.2 Edge weighting

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

Concepts and methods from graph theory are used to model and achieve this. 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 on the local transitivity of the vertices connected by the edge[5]) and a "size" factor (based on the number of 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.

---

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

[6] See supplementary material, section TODO.

2.2.3 Cohesion

The cohesion of a layer is determined from the number of fragments and connection relations 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 layers. This constraint is justified by 1) the need to use this computation to test an explicit hypothesis, and 2) the fact that archaeological spatial units are related by elementary binary relations (above to, below to, included in, etc.). (This does not prevent to study more than two spatial units by repeating the procedure for each related pairs of spatial units.) The cohesion value of each layer is given by:

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

with V_{layer} the number of vertices in the layer, V the total number of fragments in the two layers considered, W_{layer} the sum of the edge weights within the layer and W the sum of all the edge weights. Note that the relative size of each layer is also included as a factor in this computation. Considering a pair of layers, their respective cohesion values range between 0 and 1 and their sum is always equal to 1. Values toward 1 reflect higher cohesion and values toward 0 correspond to lower cohesion. E.g. 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

Considering two layers, their admixture value is equal to the “cohesion” value of a virtual third layer containing the fragments and relations at the intersection of the two layers. Therefore, it is simply computed as:

$$\text{admixture (layer1, layer2)} = 1 - (\text{cohesion}_{\text{layer1}} + \text{cohesion}_{\text{layer2}})$$

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

2.2.5 Alternative measures to compare

In this section, the TSAR method 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 on the total number of relations within the layers. I discussed the limitations of this method in the introduction.

Modularity Modularity is defined as a “quality” measure for graph partition (Newman 2006, Clauset et al. 2004). Given a graph with two groups of nodes, the 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 that the graph under study. Many methods to detect "communities" in graphs (classes of nodes with dense relations) work by optimizing modularity.

In our archaeological context, a partition is given by the association (inclusion) between fragments (nodes) and layers (node attribute). Modularity might appear as a relevant method to evaluate it. However, two reasons prevent this. First, modularity is known to have a low sensibility for small nodes groups, a flaw called "resolution limit" (Fortunato and Barthélemy 2007): it turns out that 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 cannot be connected. Modularity must therefore be abandoned in the context of the analysis of archaeological fragmentation. However, it will be included in the analysis for comparison purposes.

Topological admixture (edge betweenness centrality) A variant of topological admixture, where the edges are weighted by edge betweenness centrality, is also include. It has the advantage of relying on a previously defined and well-know 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, exploratory analyses showed that the basic behaviour of the edge betweenness give results inverse to our premise for archaeological interpretation (i.e. a connection between fragments densely connected to the fragments of their layer must receive a low weight)[7]. This can be overcome by adjusting the method without, however, resolving a second issue, namely the assignment of 0 to connections between two single fragments. Given that pairs of connecting isolated fragments are frequently observed in archaeology, this is a non-trivial issue.

2.3 Simulator

To test the different methods and to generate data to compare with empirical observations, a simulator has been designed and implemented in R language (R Core Team 2020). It can be set with several parameters (Table 2).

2.3.1 Algorithm

The TSAR simulator implements a model of archaeological fragmentation in which the topology of the relations is considered. An initial object is first broken into two fragments (Figure 5). 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 5, 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

---

[7] See supplementary material, section TODO.

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 layers, 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
aggregation factor	numerical [0;1]	when applying fragmentation, increase the likelihood of objects with more fragments to be selected
initial layers	integer [1;2]	number of initial hypothetical layers
planar	Boolean	whether generating only planar graphs

Table 2: Parameters of the TSAR simulator.

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, choose by the user, or simulated using multiple values[8].

The TSAR simulator implements this approach, using this general procedure:

1. generate n initial objects
2. select an object or a fragment
3. break it into two fragments
4. return to [1] while the number of fragments and/or number of connection relationships is not reached.

In addition, two different initial conditions can be set up: before applying the fragmentation procedure, the initial objects can be either placed in the same layer or in two different layers (Figure 6).

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 generated[9]. In summary, results are 100% accurate for the number of objects, fragments, and connection relationships. However, since these numbers must be integers, rounding of numbers is applied, leading to slight inaccuracies for the balance (median inaccuracy = -0.04±0.08 on a scale from 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

---

[8] Different 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 bellow, adjacent or non-adjacent, close or distant, layers (Rowlett and Robbins 1982, p. 79, Brantingham et al. 2007, Caron et al. 2011).

[9] See supplementary material, section TODO.

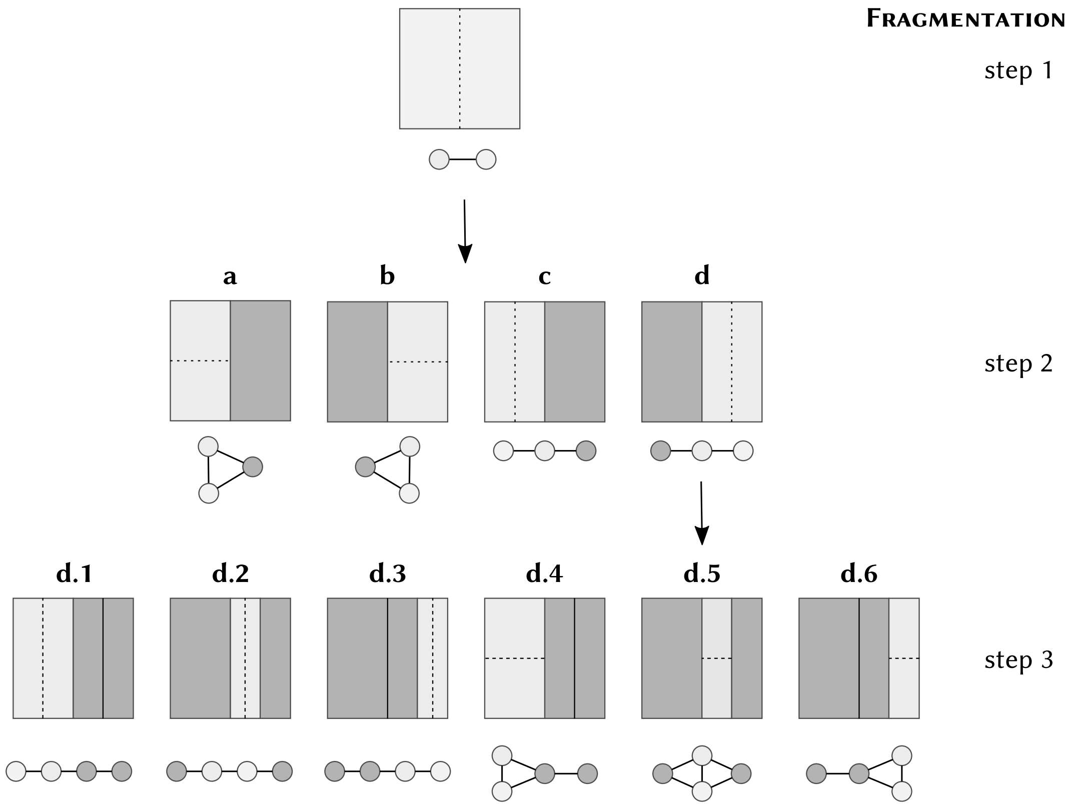

Figure 5: Model of archaeological fragmentation. Unchanged fragments are colored in dark grey. The new border is represented by dashed lines. In step (1) an object is firstly fragmented into two fragments. Four different configurations are possible in step (2) (a, b, c, d). In step (3), each configuration allows six other possibilities, 24 in total (only those related to the configuration (d) are represented here).

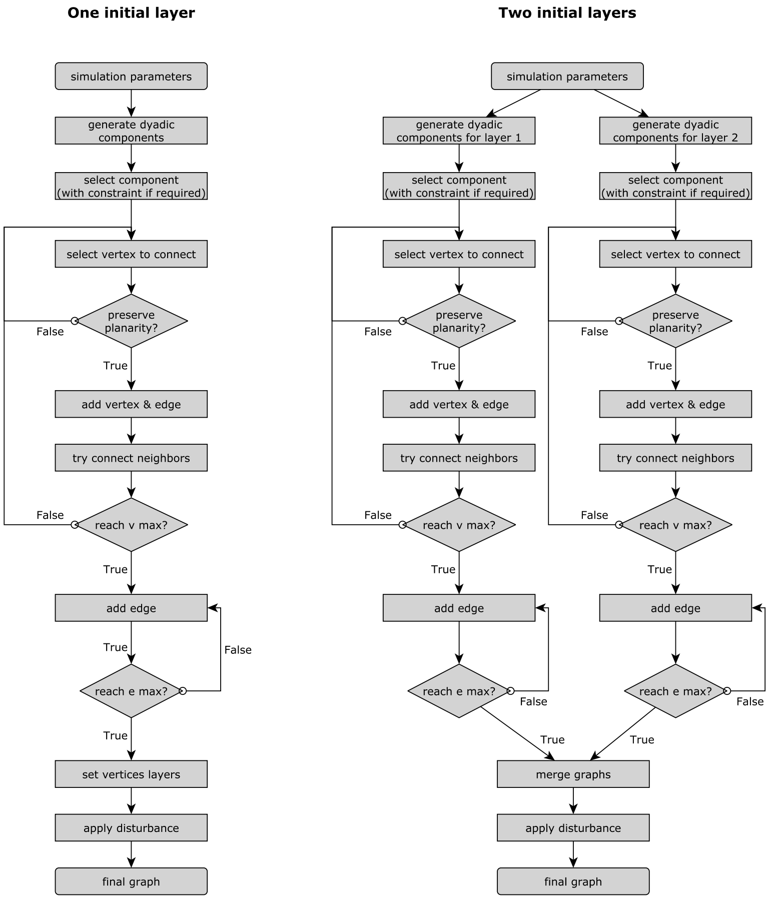

Figure 6: Flowchart of the TSAR simulation function for the two possible initial conditions: one layer (left) and two layers (right).

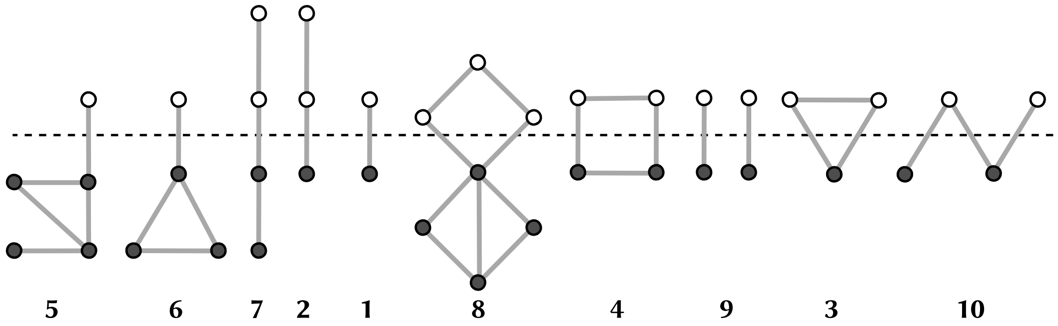

Figure 7: Theoretical examples. The colour of the nodes gives the layer of the fragment.

graphs does not affect their other parameters) and the planarity constraint has no side effects.

2.4 Data

Three data-sets are used in this study: a set of theoretical small fragmentation graphs, the connection relationships between pottery sherds founded in the Liang Abu rock shelter, the data generated with the simulator.

2.4.1 Theoretical examples

A set of 8 simplistic fragmentation graphs was defined where connected fragments are located in two layers (Figure 7). These examples were used for two purposes: to test and illustrate the different methods, and to collect and reflect archaeologists' intuitive estimation of the admixture between the two layers in each case. Using consensus modelling[10] archaeologists were asked to sort the eight graphs from the case where the layers are more distinguished (less mixed) to the case where they are less distinct (more mixed): 22 respondents were surveyed, generating 28 sorting in total, corresponding to 23 different solutions.

2.4.2 Liang Abu data set

The real world data set 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 in layer 2, raising issues specific to shallowly buried sites (Surovell et al. 2005). Two 14^c datings on charcoal from layer 2 gave coherent 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 although layers 1 and 2 have different sediments, the first being yellowish silt sediment and the second a gravel line mixed with dark brown silt sediment. This raises an interesting stratigraphical problem: is the distinction between layer 1 and 2 reliable and how to interpret it in the site

---

[10] See Whittaker et al. 1998 for a similar approach applied to pottery typology and Gnaden and Holdaway 2000 for a study of variation in stone artefacts recording.

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

Table 3: Liang Abu: number of fragments, connection relationships, and maximal number of initial objects by layer.

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

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

formation process? The study of the connection relationships between fragments contributes answering it. The data are summarised in Table 3 and Table 4.

3 Results and discussion

Results are presented in three sections, respectively 1) demonstrating the relevance and reliability of the TSAR method, 2) comparing it to alternative approaches, 3) showing its application to site formation process analysis.

3.1 Relevance and reliability of the topological method

3.1.1 Consensus analysis supports the topological method

Comparing a method result with intuitive archaeologists reasoning is a first way to assess the relevance of a method. The comparison of archaeologists' sorting and the sorting obtained by topological admixture and the three alternative measures[11] show that: 1) archaeologists gave very variable answers, which form one main and two secondary clusters; 2) topological admixture is part of the main cluster of archaeologists' answers; 3) alternative methods are grouped into a specific cluster, except modularity which is isolated (Figure 8). Consequently, the relevance of topological admixture is supported by its better fit with archaeologists' intuition.

3.1.2 Benchmark validates topological cohesion

Simulated graphs were used to benchmark the cohesion measure and test whether it reflects adequately both the effects of the relative size of two layers and of the fragments movements. Sets of graphs were generated using different pairs of values for the "balance" and "disturbance" parameters. Cohesion values and admix-

---

[11] All analysis were made using R, the *igraph* package (Csárdi and Nepusz 2006) and the *archeofrag* package (Plutniak 2020). See supplementary material.

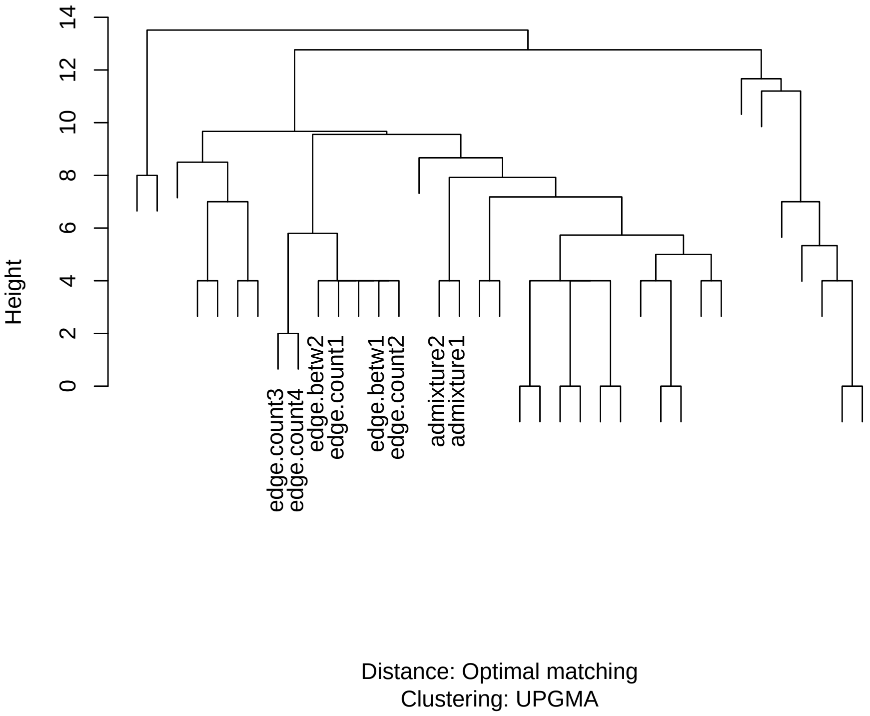

Figure 8: Clustering of the distances between the ordering of 8 theoretical examples (1 to 8, Figure 7) by archaeologists and by the four methods (topological admixture, topological admixture based on edge betweenness centrality, edge count, and inversed modularity). N.B. all the different solutions potentially generated by the methods are included and numbered.

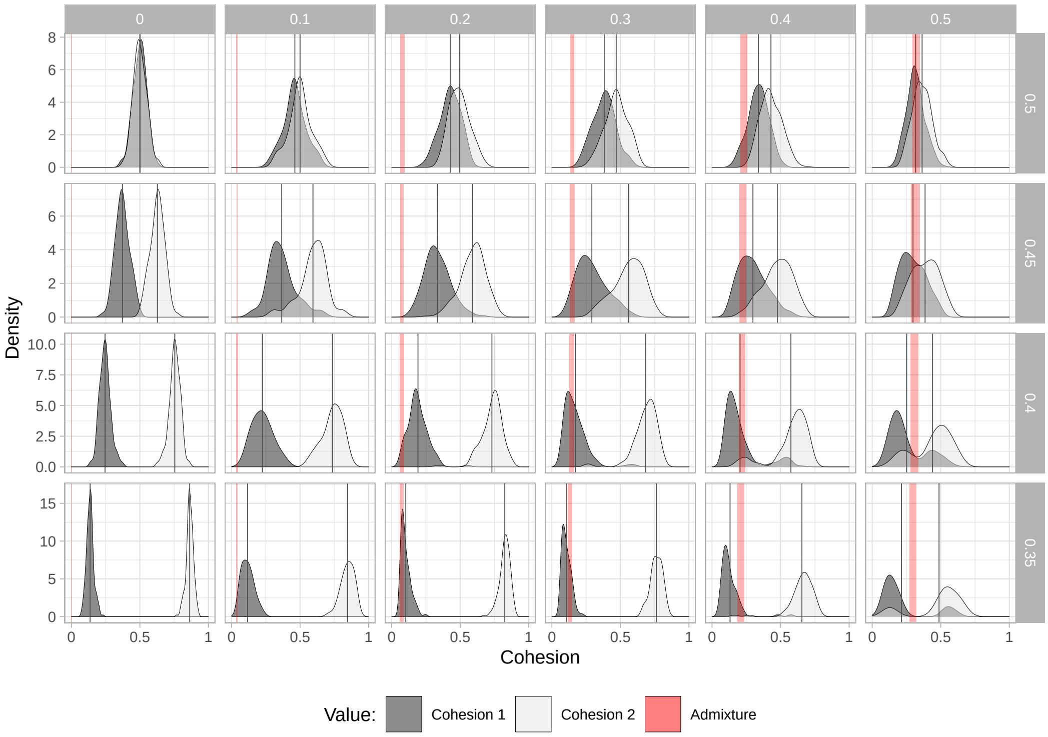

Figure 9: Density of the cohesion values for layers 1 and 2 of simulated graphs (30 objects, 100 fragments) for different balance (rows) and disturbance (columns) parameters (6000 graphs). Reading: the upper left case, for example, corresponds to two layers of similar size without inter-layer relationships. The vertical lines give the mean cohesion values. Interpretation: the relative size of the two layers is reflected by the distance between the cohesion values of the two layers. Admixture is reflected by the positions of the cohesion values: more they are left, more the layers are mixed. The red shades highlight the interquartile range of the admixture values.

ture of layers were measured on the simulated graphs (Figure 9). As evidenced by the results, similar admixture values can correspond to different proportions of the layer sizes. This confirms the need to distinguish between size proportion and fragment movements in studying the relationship between two layers. The TSAR methods fulfil this requirement as expected.

3.2 Comparison of methods

3.2.1 Ordering 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. Topological admixture and the three other methods were first applied to the theoretical examples (Figure 7, where the examples are sorted according to their topo-

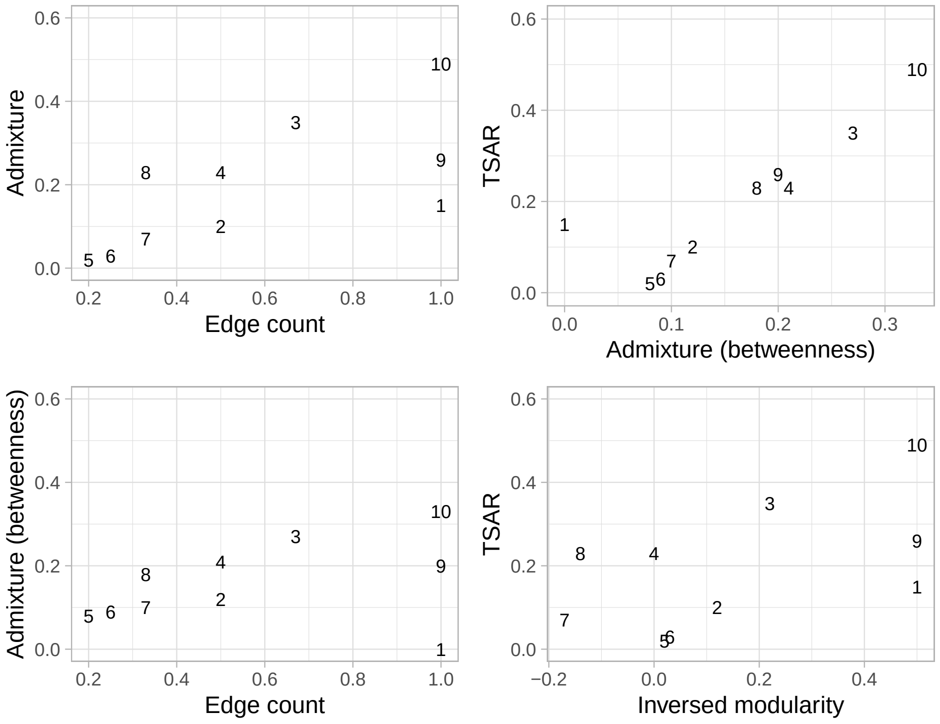

Figure 10: Comparison of the results of the four methods applied to the theoretical examples. The numbers refer to the labels of the graphs in Figure 7.

logical admixture values). Results confirm the irrelevance of modularity[12], which reports clusters only in two cases (7 and 8). Results of the edge count method show a coarse correlation with admixture results. However, edge count does not distinguish between some sets of cases, especially between 1, 9, and 10, which have clear different implications for archaeological interpretation. Edge betweenness-based admixture results are similar to admixture results, except that it assigns a null admixture to the pair of fragment located in two different layers (case 1). This is not satisfying especially because such isolated pairs of connected fragments are very frequent in archaeological observation. With topological admixture, these pairs are ranked in the middle of the series and all the cases are distinguished. There is one exception, namely configurations 4 and 8, which interestingly received the same value although an intuitive evaluation would have been unable to lead to this conclusion given the complex structure of the case 8.

3.2.2 Numerical difference between the methods

Simulated data are used to study into more details the difference between the results of the four methods. Absolute numerical difference is of importance since the archaeological interpretation relies on these numerical values ranging between 0 and 1 (if a method gives an admixture of 0.45 and a different method gives 0.55,

---

[12] To turn modularity, a measure of distinction, into a measure of mixture, inverse values were used.

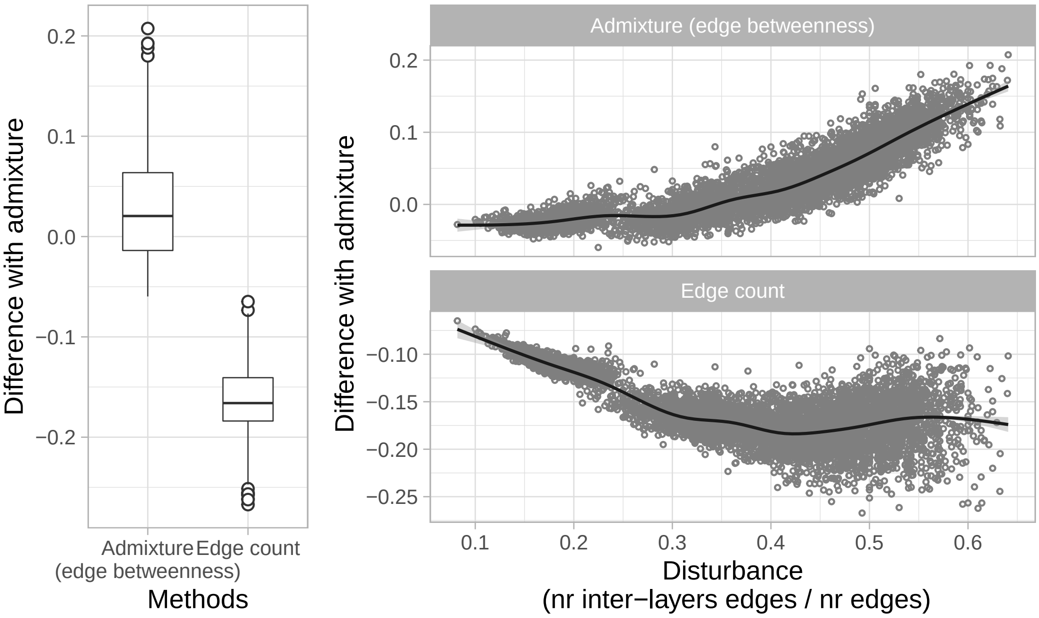

Figure 11: Difference of the results from the two methods with the admixture value observed on 1200 different artificial graphs (4 replications each) by method (left) and as a function of the disturbance observed on these graphs (right).

this difference of 0.1 might change the archaeological conclusion about the two layers considered). Non-trivial differences are observed, ranging from -0.27 to 0.21, between the results of the edge count method and edge betweenness-based admixture, compared to the results of the admixture method (Figure 11). The edge betweenness-based admixture values are higher, especially when the disturbance (proportion of inter-layers relations) increases. To the contrary, edge count results are lower and stabilised for higher disturbance values.

⁂

These results demonstrate 1) some irrelevant results generated by the alternative methods, and 2) that topological admixture is more sensible to the diverse archaeological situations, generating significantly different values.

3.3 Testing formation process hypotheses at Liang Abu

3.3.1 Layer evaluation at Liang Abu

The four methods were applied to the layers 0, 1, and 2 of Liang Abu (Table 5). By comparing layer 0 (the actual surface) and 1, layer 1 appears much more cohesive, and the admixture value is equal to 0, as expected since there is no connection relationship between them. In the case of layers 1 and 2, layer 2 has a slightly higher cohesion value and the two layers have a low admixture, 0.01. Interestingly, alternative methods gave higher admixture values. Note that modularity adequately suggests an actual distinction between both pairs of layers, with a distinction between layers 0 and 1 weaker than between layers 1 and 2. However, it must be reminded that modularity is based on assumptions irrelevant in this archaeological context.

Comparing the results from different pairs of layers from the same site is a way to give a meaning to these numbers, making them useful for archaeological interpretation. However, this approach is limited by the quantity of data available from a given site, or possibly from few other sites. Using simulated data will overcome this limitation and refine the sense of these values.

	Layers 0 & 1	Layers 1 & 2
Objects	11.00	28.00
Fragments	29.00	72.00
Cohesion layer 0	0.09	–
Cohesion layer 1	0.91	0.40
Cohesion layer 2	–	0.59
Admixture	0.00	0.01
Admixture (betweenness)	0.00	0.02
Edge count	0.00	0.06
Modularity	0.30	0.41

Table 5: 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 unique objects the sherds come from, and the cohesion, admixture and alternatives methods values.

3.3.2 Hypotheses testing

Given two distinguished layers with a non-null degree of admixture, two hypotheses about the distinction between these layers can be considered: 1) their distinction is due to multiple movements of fragments which happened in a single initial layer (e.g. post-depositional sorting process); 2) their distinction reflects the location of the fragments into two different initial layers. Each hypothesis has different consequences for archaeological interpretation. In the first case, the distinction between the two layers must be abandoned and therefore the idea of their admixture. In the second case, the distinction between the two layers must be conserved and the admixture is explained by movements of fragments.

These hypotheses are tested in the case of Liang Abu, using the properties of the graph for layers 1 and 2 as parameters for the TSAR simulator. The fragmentation process is run with two different initial conditions, namely assuming or not a boundary between two sets of objects to be fragmented (Figure 6). The properties of the graphs generated with each hypothesis are then compared to the empirical values to identify the most likely hypothesis.

Significant differences between the two hypotheses are observed about edge weights sum, balance, and cohesion values, and not for the edge count, disturbance, and admixture parameters[13] (Figure 12). About edge weight sums, Liang Abu empirical value is closer to the results generated using hypothesis 2. About balance, the empirical value matches better with the results from hypothesis 1. About disturbance, the empirical value is interestingly much lower than those

---

[13] Results supported by Mann-Whitney tests. See supplementary material, section TODO.

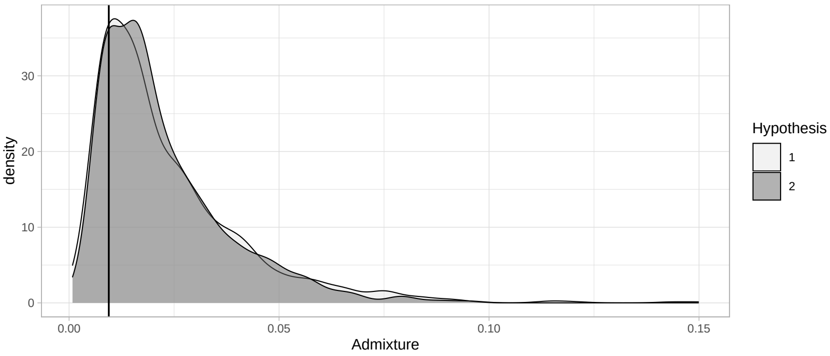

Figure 12: Admixture observed on graphs generated using Liang Abu layers 1 and 2 parameters for two formation process hypotheses (1000 replications each), compared to the admixture observed at Liang Abu (vertical bar).

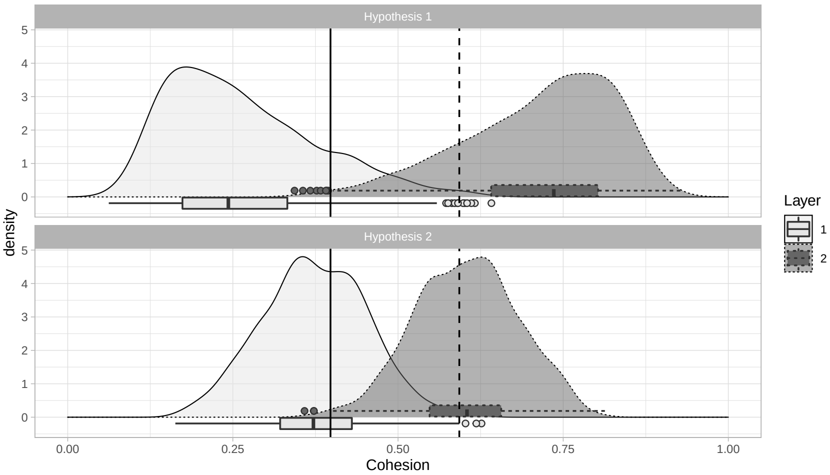

Figure 13: Cohesion values observed on graphs generated using Liang Abu layers 1 and 2 parameters for two formation process hypotheses (1000 replications each), compared to the cohesion values observed at Liang Abu (vertical bar).

from both hypotheses, which do not significantly differ. Finally, the better fitness of hypothesis 2 is better evidenced by cohesion values. Liang Abu empirical values are out of the interquartile ranges of the results simulated using hypothesis 1. To the contrary, hypothesis 2 generates result fitting perfectly with empirical values (Figure 13).

In conclusion, results support a scenario where layers 1 and 2 were initially two independent layers. The analysis of pottery fragmentation and refitting therefore validates the distinction between these layers. From a small data set, this case study illustrates how the TSAR method and simulation can enhance the analysis of fragmentation in archaeological contexts. This approach should be applied and tested on larger datasets in future work.

4 Conclusion

This paper presented a renewed framework for "refitting" analysis in archaeology. Using graph theory to model the topological relations between fragments which were formerly parts of the same original objects enables to define new measurements to assess 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 clarification and innovation. This led, in particular, to redefine "refitting" as *connection*, to give a broad definition of *layer*, and to model the *fragmentation process*.

However, more 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) applications on larger data sets, 2) weighting cohesion and admixture with morphometric values (e.g. sherd size for pottery); and 3) using the topological properties of connected fragments 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 was implemented in the *Archeofrag* R package[14]. It is complemented by a *Shiny* application[15], making easier its demonstration and dissemination.

Acknowledgements

I thank all the participants to the survey, and Joséphine Caro, Julian Laabs, Claire Manen, Luce Prignano, Oliver Nakoinz, and Émilien Schultz for their valuable comments at different steps 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).

Disclosure statement

No potential conflict of interest was reported by the author.

---

[14] Available on CRAN at `https://cran.r-project.org/package=archeofrag`.

[15] `https://analytics.huma-num.fr/Sebastien.Plutniak/archeofrag`.

Supplementary material

- .Rnw file

References

Barthès, Pascale [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.

Bollong, Charles A. [1994], "Analysis of Site Stratigraphy and Formation Processes using Patterns of Pottery Sherd Dispersion," *Journal of Field Archaeology*, 21, 1, pp. 15-28, DOI: 10.1179/009346994791549254.

Bordes, Jean-Guillaume [2000], "La séquence aurignacienne de Caminade revisitée : l'apport des raccords d'intérêt stratigraphique," *Paléo. Revue d'archéologie préhistorique*, 12, 1, pp. 387-407, DOI: 10.3406/pal.2000.1611.

Brantingham, P. Jeffrey, Todd A. Surovell, and Nicole M. Waguespack [2007], "Modeling Post-depositional Mixing of Archaeological Deposits," *Journal of Anthropological Archaeology*, 26, 4, pp. 517-540, DOI: 10.1016/j.jaa.2007.08.003.

Cahen, Daniel and Jan Moeyersons [1977], "Subsurface Movements of Stone Artefacts and their Implications for the Prehistory of Central Africa," *Nature*, 266, 5605, pp. 812-815, DOI: 10.1038/266812a0.

Caron, François, Francesco d'Errico, Pierre Del Moral, Frédéric Santos, and João Zilhão [2011], "The Reality of Neandertal Symbolic Behavior at the Grotte du Renne, Arcy-sur-Cure, France," *Plos One*, 6, 6, pp. 1-11, DOI: 10.1371/journal.pone.0021545.

Clauset, Aaron, Mark E. J. Newman, and Cristopher Moore [2004], "Finding Community Structure in Very Large Networks," *Physical Review E*, 70 [6 2004], p. 066111, DOI: 10.1103/PhysRevE.70.066111.

Csárdi, Gábor and Tamás Nepusz [2006], "The igraph Software Package for Complex Network Research," *InterJournal*, 1695, 5, pp. 1-9, http://igraph.org.

Cziesla, Erwin [1990], "On Refitting of Stone Artefacts. International Symposium on Refitting Stone Artefacts," in *The Big Puzzle*, Monrepos, Sept. 6-9, 1987, ed. by Erwin Cziesla, Sabine Eickhoff, Nico Arts, and Doris Winter, Studies in Modern Archaeology, 1, Bonn: Holos, pp. 9-44.

Cziesla, Erwin, Sabine Eickhoff, Nico Arts, and Doris 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, David, Arian Maleki, Morteza Shahram, Inam Rahman, and Victoria Stodden [2008], *15 Years of Reproducible Research in Computational Harmonic Analysis*, http://www-stat.stanford.edu/~wavelab/Wavelab_850/wavelab.pdf.

Fortunato, Santo and Marc Barthélemy [2007], "Resolution Limit in Community Detection," *Proceedings of the National Academy of Sciences*, 104, 1, pp. 36-41, DOI: 10.1073/pnas.0605965104.

Girvan, Michelle and Mark E. J. Newman [2002], "Community Structure in Social and Biological Networks," *Proceedings of the National Academy of Sciences*, 99, 12, pp. 7821-7826, DOI: 10.1073/pnas.122653799.

Gnaden, Denis and Simon Holdaway [2000], "Understanding Observer Variation When Recording Stone Artifacts," *American Antiquity*, 65, 4, pp. 739-747, DOI: 10.2307/2694425.

Gray, Lawrence [2003], "A Mathematician looks at Wolfram's New Kind of Science," *Notices of the American Mathematical Society*, 50, 2, pp. 200-211.

Harris, Edward C. [1979], *Principles of Archaeological Stratigraphy*, London: Academic Press, xiii–136 p.

Hofman, Jack L. [1986], "Vertical Movement of Artifacts in Alluvial and Stratified Deposits," *Current Anthropology*, 27, 2, pp. 163-171, DOI: 10.1086/203414.

Hofman, Jack L. [1992], "Piecing Together the Past: an Introduction to Refitting," in *Piecing Together the Past. Applications of Refitting Studies in Archaeology*, ed. by Jack L. Hofman and James Gordon Enloe, British Archaeological Reports International Series, 578, Oxford: Tempus Reparatum, pp. 1-20.

Hofman, Jack L. and James Gordon Enloe (eds.) [1992], *Piecing Together the Past. Applications of Refitting Studies in Archaeology*, British Archaeological Reports International Series, 578, Oxford: Tempus Reparatum, 321 pp.

Johnson, Donald Lee [1989], "Subsurface Stone Lines, Stone Zones, Artifact-Manuport Layers, and Biomantles Produced by Bioturbation via Pocket Gophers (Thomomys Bottae)," *American Antiquity*, 54, 2, pp. 370-389, DOI: 10.2307/281712.

Leisch, Friedrich, Manuel Eugster, and Torsten Hothorn [2011], "Executable Papers for the R Community: The R2 Platform for Reproducible Research," *Procedia Computer Science*, 4, pp. 618-626, DOI: 10.1016/j.procs.2011.04.065.

López-Ortega, Esther, Xosé Pedro Rodríguez, and Manuel Vaquero [2011], "Lithic Refitting and Movement Connections: the NW area of level TD10-1 at the Gran Dolina site (Sierra de Atapuerca, Burgos, Spain)," *Journal of Archaeological Science*, 38, 11, pp. 3112-3121, DOI: 10.1016/j.jas.2011.07.011.

Lyman, R. Lee and Michael J. O'Brien [1999], "Americanist Stratigraphic Excavation and the Measurement of Culture Change," *Journal of Archaeological Method and Theory*, 6, 1, pp. 55-108, DOI: 10.1023/A:1022976726447.

May, Keith [2020], "The Matrix: Connecting Time and Space in Archaeological Stratigraphic Records and Archives," *Internet Archaeology*, 55, DOI: 10.11141/ia.55.8.

Michel, Robert [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.

Moberg, Carl-Axel [1971], "Mathematics in the Archaeological and Historical Sciences," in *Archaeological Context and Mathematical Methods. Proceedings of the Anglo-Romanian Conference*, Mamaia, Sept. 16-24, 1970, ed. by Frank Roy Hodson, David George Kendall, and Petre Tăutu, Edinburgh: Edinburgh University Press, pp. 551-565.

Morin, Eugène, Tsenka Tsanova, Nicolay Sirakov, William Rendu, Jean-Baptiste Mallye, and François Lévêque [2005], "Bone Refits in Stratified Deposits: Testing the Chronological Grain at Saint-Césaire," *Journal of Archaeological Science*, 32, pp. 1083-1098, DOI: 10.1016/j.jas.2005.02.009.

Myers, Oliver Humphrys [1958], "'Abka Re-excavated," *Kush*, 6, pp. 131-143.

Newman, Mark E. J. [2006], "Modularity and Community Structure in Networks," *Proceedings of the National Academy of Sciences*, 103, 23, pp. 8577-8582, DOI: 10.1073/pnas.0601602103.

O'Brien, Michael J. and R. Lee Lyman [2002], "The Epistemological Nature of Archaeological Units," *Anthropological Theory*, 2, 1, pp. 37-56, DOI: 10.1177/146349960200200 1287.

Ore, Øysten [1962], *Theory of Graphs*, Providence: American Mathematical Society, 270 pp.

Plutniak, Sébastien [2020], *Archeofrag: an R package for refitting and stratigraphic analysis in archaeology*, DOI: 10.5281/zenodo.4271900, https://CRAN.R-project.org/package=archeofrag.

Plutniak, Sébastien, Astolfo Araujo, Simon Puaud, Jean-Georges Ferrié, Adhi Agus Oktaviana, Bambang Sugiyanto, Jean-Michel Chazine, and François-Xavier 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*, pp. 228-242, 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, https://www.R-project.org/.

Rowlett, Ralph M and Michael C. Robbins [1982], "Estimating Original Assemblage Content to Adjust for Post-depositional Vertical Artifact Movement," *World Archaeology*, 14, 1, pp. 73-83, DOI: 10.1080/00438243.1982.9979850.

Schaller-Åhrberg, Eva [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 Erwin Cziesla, Sabine Eickhoff, Nico Arts, and Doris Winter, Studies in Modern Archaeology, 1, Bonn: Holos, pp. 611-622.

Schurmans, Utsav [2007], "Refitting in the Old and New Worlds," in *Fitting Rocks. Lithic Refitting Examined*, ed. by Utsav Schurmans and Marc De Bie, British Archaeological Reports International series, 1596, Oxford: Archaeopress, pp. 7-23.

Schurmans, Utsav and Marc De Bie (eds.) [2007], *Fitting Rocks. Lithic Refitting Examined*, British Archaeological Reports International series, 1596, Oxford: Archaeopress.

Sisk, Matthew L. and John J. Shea [2008], "Intrasite Spatial Variation of the Omo Kibish Middle Stone Age Assemblages : Artifact Refitting and Distribution Patterns," *Journal of Human Evolution*, 55, 3, pp. 486-500, DOI: 10.1016/j.jhevol.2008.05.016.

Surovell, Todd A., Nicole M. Waguespack, James H. Mayer, Marcel Kornfeld, and George C. Frison [2005], "Shallow Site Archaeology: Artifact Dispersal, Stratigraphy, and Radiocarbon Dating at the Barger Gulch Locality B Folsom Site, Middle Park, Colorado," *Geoarchaeology*, 20, 6, pp. 627-649, DOI: 10.1002/gea.20071.

Villa, Paola [1982], "Conjoinable Pieces and Site Formation Processes," *American Antiquity*, 47, 2, pp. 276-290, DOI: 10.2307/279901.

Villa, Paola and Jean Courtin [1983], "The Interpretation of Stratified Sites: a View from Underground," *Journal of Archaeological Science*, 10, 3, pp. 267-281, DOI: 10.1016/0305-4403(83)90011-0.

Wasserman, Stanley and Katherin Faust [1994], *Social Network Analysis. Methods and Applications*, Cambridge: Cambridge University Press, xxxi–825 p.

Whittaker, John C., Douglas Caulkins, and Kathryn A. Kamp [1998], "Evaluating Consistency in Typology and Classification," *Journal of Archaeological Method and Theory*, 5, 2, pp. 129-164, DOI: 10.1007/BF02427967.

Wood, W. Raymond and Donald Lee 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 Michael B. Schiffer, pp. 539-605, DOI: 10.1016/B978-0-12-624180-8.50016-0.

Ziesaire, Pierre [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 Erwin Cziesla, Sabine Eickhoff, Nico Arts, and Doris Winter, Studies in Modern Archaeology, 1, Bonn: Holos, pp. 253-261.