



HAL
open science

Four Problems for Archaeological Refitting Studies. Discussion from the Taï Site and its Neolithic Pottery Material (France)

Sébastien Plutniak, Joséphine Caro, Claire Manen

► **To cite this version:**

Sébastien Plutniak, Joséphine Caro, Claire Manen. Four Problems for Archaeological Refitting Studies. Discussion from the Taï Site and its Neolithic Pottery Material (France). Anna Sörman; Astrid Noterman; Markus Fjellström. Broken Bodies, Places and Objects. New Perspectives on Fragmentation in Archaeology, Routledge, pp.124-142, 2023, 978-1-032-39499-2. 10.4324/9781003350026-10 . hal-04355706

HAL Id: hal-04355706

<https://hal.science/hal-04355706>

Submitted on 20 Dec 2023

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.



Distributed under a Creative Commons Attribution 4.0 International License

Four Problems for Archaeological Refitting Studies. Discussion from the Taï Site and its Neolithic Pottery Material (France)

Sébastien Plutniak¹ , Joséphine Caro² , Claire Manen² 

¹CITERES laboratory, Tours, France

²TRACES laboratory, Toulouse, France

[Author version of: Sébastien Plutniak et al. [2023], “Four Problems for Archaeological Fragmentation Studies. Discussion and Application to the Taï Cave’s Neolithic Pottery Material (France),” in *Broken Bodies, Places and Objects. New Perspectives on Fragmentation in Archaeology*, ed. by Anna Sörman et al., Routledge, DOI: [10.4324/9781003350026-10](https://doi.org/10.4324/9781003350026-10).]

Contents

1	Introduction	
1.1	Fragmented objects and uncertain archaeological location	
1.2	Data: pottery material from the Taï site	
2	Problem 1: Using similarity relationships between fragments generates bias	
2.1	Matching sets count and intuitive interpretation	
2.2	Counting and non-supervised ordering of matching relationships	
3	Problem 2: Counting connection relationships can be misleading	
3.1	Limits of count-based approaches to connection relationships	
3.2	How “strong” is a connection relationship?	
3.3	Admixture and cohesion of the Early Neolithic layers at the cave Taï	
4	Problem 3: Reference data to compare sets of refitting fragments are limited	
4.1	The limits of archaeological observation	
4.2	Data publication, experimentation, and simulation as answers	
5	Problem 4: Challenges in simulating archaeological fragmentation	
6	Conclusion	

Abstract

Determining spatial units (e.g., layers, areas) and location of archaeological remains within these units is fundamental in archaeology. However, the

fuzzy boundaries of spatial units (e.g., sediments) and the uncertainty of objects' past location (moved by possible post-depositional perturbations) are obstacles. Refitting analysis of remains are long used to reconstruct stratigraphic sequences, evaluate the impact of taphonomic processes, and detect intentional behaviour. However, four pending problems are identified and possible solutions presented. 1) Common origin is often determined from similarities in fragments. We show the inherent biases of this approach, present a new method, while arguing that physical refits only should be used. 2) Studies often count the number of refits, although this can lead to not distinguish in ambiguous cases of admixture between spatial units. The Topological Study of Archaeological Refitting method (TSAR) is applied and extended to include fragments' morphometry and distances between their locations. 3) Determining changes and fragmentation in a site formation process is limited by the impossibility to observe its past states. Three solutions are suggested (publishing refitting data, using experimental data, and simulation), fostering their collective development. 4) Current limits in computer-based simulation of fragmentation processes are presented, envisioning perspectives from high-performance computing.

Keywords: refitting, fragmentation, network analysis, post-depositional processes, stratigraphy, taphonomy

1 Introduction

1.1 Fragmented objects and uncertain archaeological location

The location of archaeological remains is of utmost importance for the interpretation of a sequence or the areas of an occupation. What is location? Formally, it is a relationship between a material entity and a region of space (Cody 2018). In archaeology, objects are usually recorded as located in spatial “archaeological units” (M. J. O'Brien and Lyman 2002) such as stratigraphic units, layers, pits, etc. These units are interpreted in order to define archaeological “contexts” and, finally, to suggest the history of a site. Nevertheless, determining the exact position of an object in a spatial unit is difficult because these units have often vague boundaries (e.g. sedimentary layers, the difference in the internal and external parts of a burial, etc.) and objects might have been moved by anthropic (J. C. Chapman and Gaydarska 2006) or non-anthropic (Wood and Johnson 1982, Schiffer 1987) post-depositional actions. Consequently, what we know about the past location(s) of a find is subject to uncertainty. Far from being a neutral observation of “raw data”, spatial units are one of the results of the archaeological process, the outcome of an archaeological investigation. Before any other consideration about contexts or intentional behaviours (Garrow 2012), field archaeologists aim at:

- determining the limits of archaeological units, and
- determining the past location(s) of objects (i.e., the –possibly changing– inclusion of objects in archaeological units) to secure their association with archaeological units.

Table 1: Four types of methods in refitting studies according to the type of relationships processed and to what is quantified (sets of fragments or sets of relationships between fragments).

		Count of:	
		sets of fragments	relationships
Types of relationship:	similarity	(1)	(2)
	connection	(3)	(4)

These aims are crucial since they form the basis on which more general statements are built. In this paper, we discuss how studying the refitting fragments of objects contribute to answer these questions.

A common ground in refitting studies in archaeology is to determine whether fragments have been parts of the same original object. Multiple methods were proposed in this regard, e.g., distinguishing between three (Cziesla 1990) or even seven (López-Ortega et al. 2011) types of lithic refits, between six types of pottery refits (Bollong 1994), or calculating the likelihood that non-refitting fragments of similar appearance are from the same object (Blanco-González and J. Chapman 2014). These methods can be categorised according to whether 1) they are based on similarity relationships (between some properties of the fragments, e.g., decoration, colour, etc.) or connection relationships (physical refit), and 2) they imply counting the number of relationships or the number of sets of fragments linked by these relationships (Table 1).

Considering a material object fragmented into multiple pieces, at the time of archaeological observation, the fragments *are* disjointed parts of the original object, but they *were* connected parts of this object in the past. “Refitting” relationships refer to these past *connection* relationships (Plutniak 2021b). They are among the most certain elementary pieces of archaeological knowledge that can be determined, because they are either true or false, non-probabilistic, determined by the principles of solid-state physics, and do not rely on external knowledge such as typo-chronologies. Connection relationships provide two aspects of fundamental archaeological information.

1. Concerning fragments and objects: connection relationships are strong evidence that two fragments that fit together were part of the same original object.
2. Concerning the location of fragments (and the complete original object): the location of a single fragment in an archaeological spatial unit is uncertain due to (anthropic or non-anthropoc) post-depositional processes. However, the more connection relationships there are within a set of fragments from the same spatial unit, the more certain it is that these fragments were primarily deposited in this unit.

Note that, in an analysis, connection relationships can be considered as also implying similarity relationships (but the opposite logic is false).

Refitting and fragmentation studies draw on a vast literature. However, this paper addresses four unresolved problems:

1. Why methods based on similarities between the properties of fragments are unsatisfactory in this context.
2. Why methods based on counting the number of refits can be misleading.
3. How the current lack of resources to compare sets of refitted fragments limits our interpretations, and how this limitation can be overcome.
4. How an estimation and reconstruction can be made of the original state of an observed fragmented assemblage.

These are general problems that can be addressed based on various types of archaeological remains (lithic, pottery, bones, etc., with possible particularities in each of these cases) and without prior assumptions about the –intentional or non-intentional– causes of fragmentation. These problems are discussed and solutions are proposed, which are then applied to the Neolithic pottery material from the Taï site (Remoulins, France)¹.

1.2 Data: pottery material from the Taï site

The Taï Neolithic site is located about 40 km from the Mediterranean Sea, in one of the valleys deeply carved in a limestone plateau, opening onto the Remoulins plain, at the intersection of several ecosystems favourable to human life. It was excavated during a period from 2001 to 2012 (Manen 2022), enabling the study of the sediment dynamics and its topography, which in turn made it possible to reconstruct the production system of the Neolithic communities, and to discuss the functional status of each occupation.

The history of this settlement was investigated through field observations, the study of artefacts, animal and vegetal resources, paleoenvironmental data and 42 radiocarbon dates. Archaeological excavations have been carried out in three sectors: the “Cave”, the cave’s extension (South “Entrance”), and the Open-air area in front of the cave. These sectors show different degrees of preservation and the chronological periods of the site are not equally represented. The sediment dynamics are characterised by an alternating succession of anthropic and natural deposits, and by erosional phases. In the Cave, the Early Neolithic (EN) deposits are well preserved, protected by the “Entrance”. The “Entrance” area was slowly eroded and detritic deposits increased in the Cave. From the Middle Neolithic onwards, human traces are less substantial in the Cave. The unprotected layers in the Open-air sector were eroded. The Late Neolithic occupation was partially preserved due to the significant erosion of the slopes, probably under anthropic pressure, covering and preserving the archaeological layers. In this paper, we draw on the Early and Middle Neolithic pottery material, presented in more detail in a previous study (Caro, Manen, et al. 2022).

¹Supplementary data to this chapter are available at <https://doi.org/10.5281/zenodo.7540404>.

Table 2: Number of sherds (NR) and minimum number of individual vessels (MNI) in the Early Neolithic sub-phases EN₁ and EN₂ at the Taï site.

Layer	Cave		South entrance		Total	
	NR	MNI	NR	MNI	NR	MNI
EN ₁	540	315	106	73	646	388
EN ₂	595	362	288	177	883	539
Total	1135	696	394	251	1529	947

The earliest human occupations belong to the Early Neolithic (EN) and are dated to 5270–4990 cal. BC. These layers are well preserved in the Cave and in the Entrance (about 40 m²) but have been totally eroded in the Open-air sector. However, the dwellings probably extended to the Open-air area and the Cave could have been used for domestic activities and as a discard area. The EN layers were divided into two subsets (EN₁ and EN₂) based on stratigraphic arguments and the nature of the sedimentary units (silo and pit filling for EN₁ and circulation layers with hearths for EN₂). Pottery material from the Cave and the South Entrance includes 1529 sherds associated with the Epicardial techno-complex (Table 2).

After an abandonment period, the site was occupied during the Middle Neolithic (MN), from 4010 to 3700 cal. BC. The MN layers are poorly preserved in the cave, well-developed in the “Entrance” and totally eroded in the open-air sector. The lithic industry and the pottery material (N=2486 sherds, from the South Entrance) are associated with the *Chasséen* techno-complex. Domestic structures are rare but the deposits present a high density of combustion residues.

Grouping issues have been raised when interpreting this sequence, concerning the EN phase, on the one hand, and the MN phase, on the other hand. Refitting and fragmentation were used to answer these problems: the rest of this paper discusses and illustrates general methodological issues in fragmentation studies using the EN and MN case studies.

2 Problem 1: Using similarity relationships between fragments generates bias

During the excavation, six layers were distinguished in the MN phase (based on the sedimentary characteristics and the dispersal of the material), and grouped into two subgroups based on stratigraphic and pottery data: MN₁ (Layers 24, 26, 28) and MN₂ (Layers 29, 30, 31). However, the association of Layer 29 with MN₂ is ambiguous: radiocarbon data and results from the study of lithic materials suggested associating Layer 29 with MN₁. This section’s aim is to examine whether the distribution of the matching relationships between fragments supports this division or not.

The most common and intuitive approach to archaeological fragmentation in refitting studies has long been to determine and quantify sets of fragments. However, the quantification can concern either the sets of fragments (cases 1 and 3 in Table 1) or the relationships between these fragments (case 2 in Table 1). In this

section, the two approaches are presented and their limits are discussed, paying particular attention to similarity relationships.

2.1 Matching sets count and intuitive interpretation

The study of the 2486 pottery sherds from the MN phase in the South Entrance led to the identification of 37 “similarity” sets² (101 sherds; 4% of the total), as well as 71 “connection” sets (251 sherds; 10% of the total). These sets can include fragments from different layers: 77 sherds are included in interlayer “similarity” sets, and 29 sherds are included in interlayer “connection” sets. Of these 108 similarity sets and connection sets, 84 sets (78%) include only fragments from the same layers, the interpretation of this could be that there was little mixing between layers. However, the remaining 24 sets (22%) with fragments from different layers can be used to test the division between MN1 and MN2. Examining Table 3 suggests that the upper part MN1 (Layers 24, 26, 28; 16 sherds in total) has fewer interlayer relationships than the bottom part MN2 (Layers 29, 30, 31; 88 sherds in total). This supports the distinction between MN1 and MN2 and opposes the association of L29 with MN1.

Nevertheless, several problems are raised by this usual approach: 1) connection and similarity relationships have a similar analytical status, although the first type of relationship is stronger evidence of the fragments being part of a whole³ than the latter type of relationship. 2) The number of fragments included in each set is not taken into account. 3) There is no clear method, criterion or threshold to determine whether a spatial unit is well defined, slightly mixed, or too much mixed and so irrelevant, finally relying on the archaeologists’ subjective intuition and experience. Consequently, a more rigorous procedure is needed for the analysis of similarity relationships between fragments, as already stated in a previous attempt to characterise “archaeological similarity networks” (Prignano et al. 2017).

2.2 Counting and non-supervised ordering of matching relationships

In this second approach, the focus is on relationships rather than on sets of fragments. Similarity and connection relationships are regarded indiscriminately as matching relationships. Let us consider a matrix with the number and proportion of relationships between and within layers (Table 4). Reading its diagonal reveals that Layers 26 and 30 have higher proportions of internal relationships, strongly supporting their determination as spatial units.

The number of relationships can be interpreted as a measure of similarity: the more two layers are linked by relationships, the more likely it is that these two layers constitute the same archaeological spatial unit. Theoretically, a spatial unit is expected to be more similar to those near it. In the case of stratigraphic layers, a layer is expected to be more related to the layers directly above and below it. Thus, if we convert the similarity measure into a dissimilarity measure (a distance, in the mathematical sense), hierarchical cluster analysis methods can be applied.

²In this analysis “similarity” sets include only sherds with similarity relationships, whereas “connection” sets include at least one connection relationship and the related similar fragments.

³It is what philosophers call a “parthood” relation (Cody 2018).

Table 3: Number of sherds according to layers (columns), divided into interlayer sets {similarity sets} and [connection sets] (rows) in the MN phase.

Set id	MN1		MN2			
	L24	L26	L28	L29	L30	L31
[10]	2	6				
{27}	1			2		
{51}		1		1		
[66]		2			2	
{26}				1	1	
{38}				1	1	
{41}				1	1	
[63]				1	1	
{48}				2	1	
[68]				2	1	
[71]				2	2	
[41]				3	4	
{49}				1	2	
{49}				1	2	
[54]				1	3	
[56]				1	9	
[62]				3		2
{63}				1		1
[60]				1	5	2
[84]			1			1
[29]		1		1	6	1
[52]	1			4	4	
{70}	1			1	1	1
{25}					1	1
{34}					1	3
Total	5	10	1	30	46	12

Table 4: Number and proportion of similarity relationships within and between the layers of the Middle Neolithic phase (966 relationships, 72 similarity sets).

	L24	L26	L28	L29	L30	L31	L32
L24	44 (5%)						
L26	28 (3%)	261 (27%)					
L28	0 (0%)	20 (2%)	25 (3%)				
L29	9 (1%)	58 (6%)	7 (1%)	76 (8%)			
L30	5 (1%)	31 (3%)	9 (1%)	82 (8%)	120 (12%)		
L31	1 (0%)	24 (2%)	11 (1%)	26 (3%)	69 (7%)	46 (5%)	
L32	0 (0%)	4 (0%)	0 (0%)	0 (0%)	4 (0%)	4 (0%)	2 (0%)

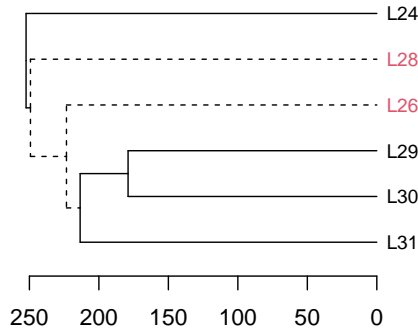


Figure 1: Hierarchical classification of the layers of the MN phase. Inverted layers are highlighted by dashed lines and red labels. Data: x , a matrix with the number of similarity relationships (Table 4); distance: $dist_{ij} = max(x) - x_{ij}$; clustering method: UPGMA.

We use the UPGMA algorithm (Unweighted Pair Group Method of Aggregation), a bottom-up method based on the average of the distances between the objects of the concerned classes at each iteration of the algorithm (Sokal and Michener 1958). Once the dendrogram is obtained, its branches are ordered alphanumerically according to their label (following the stratigraphic order of the layers). Anomalies are revealed when, despite this ordering constraint, the expected order of superposition is not observed in the result.

This method is applied to the six layers of the MN phase. In the result, all the layers are ordered as expected, except L26 and L28 which are inverted⁴ (Figure 1). Re-examining Table 3 helps to explain this anomaly: L26 has more relationships with distant layers (L29, L30, L31) than with the closer L28. Consequently, the definition of MN1 as a unit including Layers 24, 26, and 28 is not supported. This result qualifies those obtained through the intuitive examination of the spatial distribution of the similarity sets. Nevertheless, both methods confirm the association of L29 with MN2, contradicting results from radiocarbon dating and lithic analysis. In summary, the method based on the count of relationships allows to a more accurate estimation of the link between all pairs of layers, confirming the correct ordering of the layers in the sequence while detecting an anomaly (L26, which might be explained by a failure, during the excavation, in detecting a cut in L28).

This approach offers 1) an explicit method to analyse matching relationships defined by similarity and 2) a criterion to detect anomalies. Nevertheless, it suffers from the bias specific to similarity relationships: the fact that all the fragments of a matching set are, by definition, linked together, which leads to a multiplication

⁴Running the same analysis with connection or similarity relationships only leads to similar results.

of the number of relations without these being validated by direct observation. This follows a geometrical progression in which n fragments have $(n \times (n-1))/2$ relations, meaning that 2 fragments have 1 relation, 3 fragments have 3 relations, 4 fragments have 6 relations, 5 fragments have 10 relations, and 10 fragments have 45 relations, etc. A solution to overcome this issue is to only use fragments that physically fit together, i.e. connection relationships.

3 Problem 2: Counting connection relationships can be misleading

3.1 Limits of count-based approaches to connection relationships

Let us now consider using only matching sets defined by connection relationships (case 4 in Table 1). This approach is also unsatisfactory, raising two issues that have been discussed in a previous study (Plutniak 2021b). First, supposing we have two archaeological situations with identical numbers of refits between objects, and identical proportions of refits within and between spatial units. In such a case, we have demonstrated that different conclusions can be drawn in these two situations about whether the distinction between the two spatial units actually corresponds to two deposition events and about the first location of the objects when deposited. Second, submitting to archaeologists a series of examples where refitting fragments are differently distributed in two spatial units, and asking them to evaluate how mixed these units are, revealed a low degree of consensus. This demonstrated the need for a formal and explicit method to “weigh” the value of connection relationships in order to differentiate these ambiguous cases⁵. This has been achieved by supplementing the count of connection relationships by considering their relative spatial distribution (their “topology”, i.e. the “structure” of the network of connection relationships between refitted fragments)⁶. This was implemented in a method coined TSAR (Topological Study of Archaeological Refitting, Plutniak 2021b), using graph theory to model connection relationships: this section presents an extension of the TSAR method and applies it to the pottery material from the Tai EN phase.

3.2 How “strong” is a connection relationship?

As already mentioned, the more connection relationships there are between a set of fragments from the same spatial unit, the more certain it is that they were originally deposited in this spatial unit. Consequently, non-refitting fragments do not provide information in this regard and should not be considered. In the TSAR method, the weighting procedure aims to evaluate each connection relationship by giving it a number representing the “cohesive” strength of association (between

⁵Considering similarity relationships, some authors defined scales of certainty in this context (e.g., relations 3 and 4 in Bollong 1994, p. 18; Blanco-González and J. Chapman 2014, p. 250-251); however, determining degrees of certainty remains very dependent to subjective evaluation.

⁶Note that this is not an entirely new idea in this context (Gouletquer 1979, p. 26-29, Cziesla 1990, Michel 2002, p. 122-123), but an idea that has not been developed and applied.

Table 5: Interpretation of the cohesive “strength” of connection relationships as a function of the relative size of the connected fragments and the distance between their findspots.

		Distance	
		Close	Distant
Relative size	Large-Large	+	---
	Large-Small	++	--
	Small-Small	+++	-

the linked fragments and their spatial unit) provided by this relationship. By default in TSAR, the weighting is computed based on two structural properties of the graphs representing the network of connection relationships: the degree (number of relationships that are incident to a fragment) and the local clustering coefficient of the fragments⁷, as well as a factor based on the relative size of the data sets for each spatial unit (in number of fragments and relationships). This is the basis of the TSAR method.

However, other factors can be considered to determine the “strength” of connection relationships. Let us assume that:

1. The size of the objects influences the probability that they might have been moved: small objects are more likely to be transported by natural processes than large ones.
2. The distance between connected fragments when they were reflects the significance of the (intentional or unintentional) processes that disturbed the object and/or their original spatial units.

These assumptions and their interpretations are summarised in Table 5: two connected small fragments found near each other suggest a higher cohesion than two large fragments found far from each other.

The TSAR method has been expanded to include an optional *morphometric factor* (the size of the refitting fragments)⁸ and an optional *distance factor* (the distance between the places where they were found) in the weighting procedure⁹. Three variables are combined in order to translate these principles into a factor¹⁰ used to alter the values computed with the topological method only:

1. the distance between the findspots of the fragments.

⁷In graph theory, the local clustering coefficient (or transitivity) is the number of edges between the vertices within its neighbourhood divided by the number of edges that could exist between them.

⁸Note that morphometric data are used here in relation to transport processes, and not in relation to fragmentation processes.

⁹This feature has been implemented in the R package *archeofrag* since version 0.7 (Plutniak 2022a).

¹⁰See supplementary materials for a study and a calibration of this factor using numerical simulations.

2. the size of the fragments (either in terms of length, length by width, area, volume, etc.);
3. the relative sizes of the fragments (are they similar or different in size?).

These variables are normalised to the maximum values observed in the studied data set, such as:

$$f(E_{ij}) = 1 - \left(\sqrt{\frac{size_i + size_j}{max(sizes)}} \times \sqrt{\frac{\frac{size_i}{size_j}}{max(proportions)}} \times \frac{distance_{ij}}{max(distances)} \right)$$

with, for each edge $E(i, j)$ connecting fragments i and j :

- $size_i$ and $size_j$: the morphometric values for fragments i and j ;
- $distance_{ij}$: the spatial distance between fragments i and j .

and

- $sizes$: the series of sums $size_i + size_j$;
- $proportions$: the series of ratios $size_i/size_j$, with $size_i < size_j$;
- $distances$: the series of spatial distances between fragments i and j .

The morphometric-distance factor is in the range of]0;1] and is used as a multiplier of the value calculated with the topological weighting procedure. To illustrate this method, it has been applied to the Taï EN pottery material.

3.3 Admixture and cohesion of the Early Neolithic layers at the cave Taï

A graph was built to model the connection relationships within and between the EN1 and EN2 layers for the Cave sector, on the one hand, and for the South entrance, on the other hand (Table 6). The number of connection relationships is slightly similar in both sectors, but there are many more fragments and matching sets in the Cave.

When assessing the relevance of the distinction between the layers EN1 and EN2, the TSAR method provides the means to measure the internal cohesion of each spatial unit and their admixture (based on the fragments and their relationships). These values were calculated using different weighting methods for comparison:

1. the topological factor only;
2. the topological factor modified by the morphometric factor (multiplying the length of the fragments by their width);
3. the topological factor modified by the morphometric-distance factor.

Results are presented in Table 7.

Table 6: Properties of the fragmentation graphs built for the Cave and South entrance sectors (see the [Plutniak 2021b](#), Table 2, and the *archeofrag* documentation for details).

Parameters	Cave	South entrance
Number of matching sets	31	18
Number of fragments	92	63
Number of relations	78	68
Fragment balance EN ₁ /EN ₂	0.33	0.28
Object balance EN ₁ /EN ₂	0.36	0.18
Perturbation	0.08	0.02
Aggregation	0.79	0.85
Planar	yes	yes

Results In the Cave, EN₁ shows significantly higher cohesion values than EN₂, whereas this difference is less significant in the South entrance. The admixture value is also higher in the Cave than in the South entrance. These results stress the variability that might be observed within a single site, due to the different development of its sectors and variation in the sampling method. Different formation and deposition processes occurred in the two sectors. In the Cave, the EN layers are made up of numerous pits used as dumps, filled with more “disconnected” discards. In the South entrance, EN layers instead correspond to “soil layers”, that are more coherent. Interpreting the cohesion and admixture values (see [Plutniak 2021b](#), Table 1, for details about the interpretation method) supports the identification of Layer EN₂ and suggests that fragments were moved from this layer (without properly supporting the identification of a Layer EN₁ to which the fragments might have moved).

Effects of the morphometric-distance factor Using this case study, these results allow estimation of the effects of the three weighting methods on the cohesion and admixture values. Compared to the use of the topological factor only, using morphometric data tends to reduce the difference between layers EN₁ and EN₂ in regard to the degree of cohesion. However, using morphometric data and spatial distances results in values similar to those generated with the topological method only, with a slightly lower admixture value in the Cave sector whereas in the South Entrance admixture values stay very low with the three methods.

*

**

In summary, this method allows calculation of the cohesion and admixture of pairs of archaeological spatial units on a scale from 0 to 1. The resulting values contribute to interpretation of the history of the part of a site under consideration. However, this approach has limits, since 1) it applies only to pairs of spatial units, and 2) it does not include references to compare and interpret the computed values. In other words, some issues related to the comparison of sets of connected fragments remain unsolved.

Table 7: Cohesion and admixture values for EN1 and EN2 by sector and by weighting method (“x, y, z” is a shorthand for the distance factor).

Sector	Method	Cohesion		Admixture
		EN1	EN2	
Cave	topology	0.161	0.806	0.033
Cave	topology + morphometry	0.197	0.775	0.028
Cave	topology + morphometry + (x, y, z)	0.163	0.809	0.028
South entrance	topology	0.346	0.654	0.001
South entrance	topology + morphometry	0.405	0.594	0.001
South entrance	topology + morphometry + (x, y, z)	0.377	0.623	0.001

4 Problem 3: Reference data to compare sets of refitting fragments are limited

4.1 The limits of archaeological observation

Whatever the method (counting matching relationships, counting connection relationships only, the TSAR method), all generate numbers that must be interpreted to determine 1) the archaeological relevance of the definition of a spatial unit as corresponding to a deposition event, and 2) the processes (intentional or not) that caused fragmentation and displacement. However, how can these numbers be interpreted? For example, considering two layers, what would be the maximum proportion of interlayer connection relationships before considering that the two layers actually correspond to a single spatial unit? Would this threshold occur at 2%, 5%, 10%, or even more? How can this threshold be defined other than in a subjective and arbitrary manner? A method might be to determine the extent of changes that occurred in an archaeological sequence. This would involve observing its current state and comparing it with a series of past states. Travelling in time to carry out this series of observations is, unfortunately, impossible. Similarly, reference collections and experimental data sets might also be used to determine thresholds, but these types of data are far too infrequent.

A possible solution, using the TSAR method, is to compare several pairs of spatial units from the same context. It would allow us to state, for example, that layers A and B are less or more mixed than layers B and C. Such an approach was presented above concerning similarity relationships (Figure 1) and could also be applied to connection relationships, by considering the admixture of fragments as a measure of similarity between spatial units. However, this would only allow comparisons at a *local* scale (e.g., a site); the problem of comparisons at a large scale remains open. In this context, three solutions are considered in this section for future fragmentation studies: 1) the publication of refitting data sets, 2) the generation and publication of experimental fragmentation data, 3) and the use of computer-based simulations to generate unobservable situations.

4.2 Data publication, experimentation, and simulation as answers

Publishing data sets The first potential means of making data available for comparison is to promote the publication of data sets, related to a broad range of archaeological sites, periods, and types of materials (pottery, lithic, fauna, etc.). Open access and non-profit platforms are preferred for this purpose (e.g. *Zenodo*¹¹, *Open Science Framework*¹², *Nakala*¹³). For example, pottery refitting data from the Liang Abu rock shelter (*Plutniak 2021a*) and the Tai site (*Caro and Plutniak 2022*) are published in this way. Some data sets are also available on *The Digital Archaeological Record* (tDAR). In addition, tools to make these data sets discoverable on the web should be developed; in particular, indexing concepts in documentation systems. More than thirty years ago, Erwin Cziesla contributed to this aim by publishing a trilingual vocabulary (French, German, English) including 12 concepts related to refitting studies (*Cziesla 1990*, p. 36). Some actual documentation systems include similar concepts: the multilingual *PACTOLS* thesaurus for archaeology lists the concept of “refit” related to lithic material¹⁴, the *RAMEAU* thesaurus of the French National Library lists the concept of “remontage”¹⁵, and the German *iDAI.world Thesaurus* includes several related concepts, namely “*Refitting-Methoden*”, “*Zusammensetzung*”, and “*Keramikzusammensetzung*”¹⁶. These concepts should be complemented, enriched, made multilingual, or created in the documentation systems that do not yet include them¹⁷.

Experimentation A second means to assess refitting data from excavations, and to compare this with reference data, is to create experimental data sets. As already mentioned, discard and deposition experiments enable direct observation of the stages of a site formation process, which cannot be achieved at actual archaeological sites. Previous taphonomic experimental approaches were carried out at the level of site formation processes (*Schick 1986*), or at the level of an assemblage, e.g. work studying the abrasion (*Schiffer and Skibo 1989*) or salt erosion (*P. O’Brien 1990*) of pottery surfaces, or the morphometry of pottery sherds (*Rutkoski 2019*). Experimental approaches including fragmentation analysis are even rarer, e.g., refits of pottery sherds and bones (*Villa and Courtin 1983*), with focus on the weight, number, and arc of rims of pottery sherds (*Chase 1985*), on pottery (*Malloy 2019*) or bone (*Morin et al. 2021*) dispersion. A qualitative study of fragmentation patterns has also been proposed (*García Rosselló and Calvo Trias 2019*). Integrating fragmentation studies and refitting studies –both in the field and in the laboratory– is still an open challenge. This involves, not only site formation processes, but also to determine if intentional breakage can be detected from the properties of refitting sets of fragments (*Evans and Barrera Hernandez 2017*). It can be addressed using

¹¹<https://zenodo.org>.

¹²<https://osf.io>.

¹³<https://nakala.fr>.

¹⁴The permanent URI of this concept is: <https://ark.frantiq.fr/ark:/26678/pcrtSy10ihyt5s>.

¹⁵<https://catalogue.bnf.fr/ark:/12148/cb159109510>.

¹⁶Their URI are http://thesauri.dainst.org/_8d28125c, http://thesauri.dainst.org/_3ceb7211, and http://thesauri.dainst.org/_8ba55fff, respectively.

¹⁷E.g., *the Library of Congress Subject Headings*, *Wikidata*, *DBpedia*.

the TSAR descriptive methods, paying particular attention to the publication of the resulting data sets, in order to enable comparisons.

Using artificially generated data A third means to overcome the lack of empirical reference data is through simulations and artificially generated data sets. In this context, simulation has been used to estimate the size of an original set of objects based on observed fragments (Orton 1982, Felgate et al. 2013), or to model post-deposition processes (Brantingham et al. 2007, Carver 2015). A simulation-based approach to fragmentation, with particular focus on connection relationships, was developed within the framework of the TSAR method, presented in detail in previous works (Plutniak 2021b, Plutniak 2022b), and implemented in the *archeofrag* package (Plutniak 2022a). Let us consider the timeline of an archaeological assemblage, where t_0 corresponds to the state of fragments as appearing during excavation (Figure 2). Archaeologists are interested in reconstructing past states of this assemblage. Determining refitting fragments allows for a partial reconstruction of the objects of the assemblage, which represents a past state t_{-1} . Note that t_{-1} is a fictional state, i.e. there is no way to control whether the partially reconstructed objects (at t_{-1}) have existed contemporaneously (Bordes 1980). Consequently, this state is no more than a reasonable fiction, an approximation of past reality that is useful for archaeological investigation. What the archaeologist observes in the field (at t_0) involves pure spatial relationships, the states t_0 and t_{-1} are like snapshots of a process. What happened during the time between t_{-1} and t_0 , processes such as movement, cannot be directly observed. This simulator enables simulating different fragmentation scenarios from t_{-1} to t_0 . This approach has been applied to the Tai MN phase in a previous study, suggesting the scenario where MN1 and MN2 were differentiated within a single original deposition unit (Caro, Manen, et al. 2022, p. 625-630). Nevertheless, challenges in archaeological fragmentation simulations remain.

5 Problem 4: Challenges in simulating archaeological fragmentation

Reconstructing the state of an assemblage as in t_{-1} is not sufficient; archaeologists have long been interested in approximating the original composition of complete assemblages, i.e. the number of objects (Orton 1982, Felgate et al. 2013)¹⁸. Consequently, a current challenge is to address what happened in the time gap between the moment corresponding to the partial reconstruction of objects (t_{-1}) and, earlier, the “beginning” of the deposition process, i.e. the theoretical initial and original state when all the objects were complete (t_{-2} in Figure 2). Multiple scenarios of fragmentation and movement are possible. Note that, like the t_{-1} state, the t_{-2} state is a convenient fiction; there is no reason to believe that all the considered objects were complete contemporaneously at some point in the past.

Given a fragmentation model (e.g., the model implemented in the TSAR simulator), this is assumed to determine which original states at the time of t_{-2} that best

¹⁸Note that advanced computer-based modelling and simulation were also used for a near – although different – aim, automating the identification of refitting (e.g., Cooper and Qiu 2006).

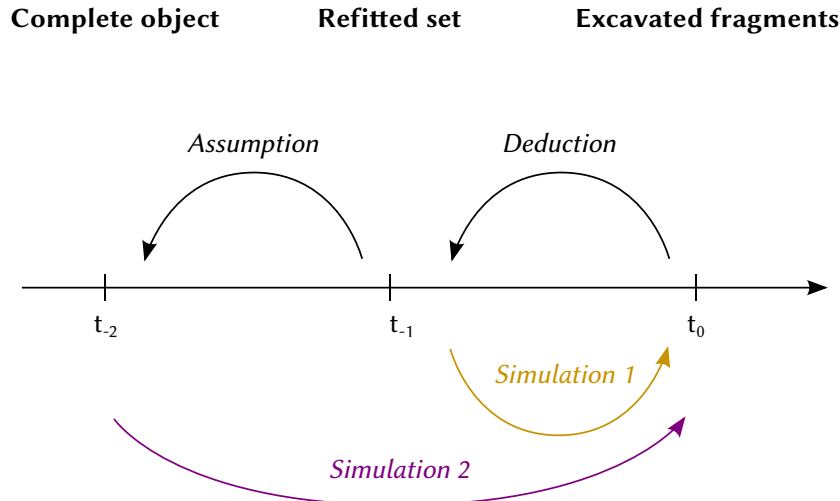


Figure 2: Timeline of an archaeological assemblage: type of reasoning applied to determine past scenarios, based on empirical observation (top) and uses of simulation (bottom).

simulate the fragmentation and movement processes that might have led to the t_0 state. This raises a serious problem since assumptions about the initial parameters of the simulation are required (e.g., the initial number of objects and spatial units, the initial distribution of objects in the spatial units, etc.); combinations of possible initial conditions are too numerous to consider. However, recent modelling approaches based on high-performance computing, such as “model exploration” methods, can help. The “Origin Search Exploration” method (OSE) is of particular interest in this context; from an output of a model it enables determination of possible combinations of the initial parameters. Conducting an OSE analysis requires, first, a definition of 1) the parameters values for an observed state of a model (e.g., the values describing the state at t_0), and 2) the ranges of possible initial values for each parameters of the model. Second, running the OSE procedure led to the determination of which combinations of values that best generate the observed state and, consequently, the most probable t_{-2} states. OSE is implemented in the *Open Mole* platform, which provides convenient access to model exploration methods¹⁹. Ongoing research is currently carried out in this direction.

6 Conclusion

This paper contributes in two ways to current and future refitting and fragmentation studies. It identifies four problems and has offered –or envisioned– solutions related to: 1) the use of similarity relationships in refitting studies, 2) count-based approaches to refits, 3) the need of reference data to compare fragmented materials, and 4) the simulation of fragmentation processes. Although (1) and (2) are commonly used in archaeology, related underrated limits and biases have been em-

¹⁹<https://openmole.org/OSE.html>.

phased, justifying new methodological developments. The use of the topological approach to refitting (the TSAR method) has been illustrated based on the case of pottery material from the Taï site. The importance of combining insights from refitting studies and from other approaches in the interpretation (direct observation of the stratigraphy, geoarchaeology, stylistic analysis, etc.) has been stressed. Problems (3) and (4) are still open and require more investigation, which might lead to further development of fragmentation studies and better understanding of formation processes at archaeological sites. Here, distinguishing between spatial units in stratigraphic analysis was particularly discussed. However, the four issues addressed cover a wider spectrum of archaeological questions, including distinguishing between forms of intentional and non-intentional breakage. This will be addressed in future studies.

References

- Blanco-González, Antonio and John Chapman [2014], "A New Method for Identifying Sherd Refits: A Case Study from the Neolithic of Northumbria, U.K.," *Journal of Field Archaeology*, 39, 3, p. 248-255, DOI: [10.1179/0093469014Z.00000000083](https://doi.org/10.1179/0093469014Z.00000000083).
- Bollong, Charles A. [1994], "Analysis of Site Stratigraphy and Formation Processes using Patterns of Pottery Sherd Dispersion," *Journal of Field Archaeology*, 21, 1, p. 15-28, DOI: [10.1179/009346994791549254](https://doi.org/10.1179/009346994791549254).
- Bordes, François [1980], "Question de contemporanéité : l'illusion des remontages," *Bulletin de la Société préhistorique française*, 77, 5, p. 132-133.
- Brantingham, P. Jeffrey, Todd A. Surovell, and Nicole M. Waguespack [2007], "Modeling Post-depositional Mixing of Archaeological Deposits," *Journal of Anthropological Archaeology*, 26, 4, p. 517-540, DOI: [10.1016/j.jaa.2007.08.003](https://doi.org/10.1016/j.jaa.2007.08.003).
- Caro, Joséphine, Claire Manen, Agathe Baux, and Sébastien Plutniak [2022], "Les productions céramiques du Néolithique ancien et moyen : approches céramo-stratigraphique, technologique et morpho-stylistique," in *Le Taï (Remoulins – Gard). Les premières sociétés agropastorales du Languedoc méditerranéen. Le Taï (Remoulins – Gard)*, ed. by Claire Manen, Toulouse: Archives d'Écologie Préhistorique, p. 616-687.
- Caro, Joséphine and Sébastien Plutniak [2022], *Refitting Neolithic Pottery Fragments from the Taï site, France*, Zenodo, DOI: [10.5281/zenodo.7408706](https://doi.org/10.5281/zenodo.7408706).
- Carver, Geoff [2015], "Modelling the Effects of Post-depositional Transformations of Artifact Assemblages using Markov Chains," in *Mathematics and Archaeology*, ed. by Juan A. Barceló and Igor Bogdanovic, Boca Raton (Fla.): CRC press, p. 293-313.
- Chapman, John C. and Bissierka Ivanovna Gaydarska [2006], *Parts and Wholes. Fragmentation in Prehistoric Context*, Oxford: Oxbow Books.
- Chase, Philip G. [1985], "Whole Vessels and Sherds: An Experimental Investigation of their Quantitative Relationships," *Journal of Field Archaeology*, 12, 2, p. 213-218, DOI: [10.1179/jfa.1985.12.2.213](https://doi.org/10.1179/jfa.1985.12.2.213).
- Cody, Gilmore [2018], "Location and Mereology," in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta, <https://plato.stanford.edu/archives/fall2018/entries/location-mereology>.
- Cooper, Judith R. and Fang Qiu [2006], "Expediting and Standardizing Stone Artifact Refitting using a Computerized Suitability Model," *Journal of Archaeological Science*, 33, 7, p. 987-998, DOI: [10.1016/j.jas.2005.11.005](https://doi.org/10.1016/j.jas.2005.11.005).

- 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, p. 9-44.
- Evans, S. and Sandra Barrera Hernandez [2017], "Sherd Shatter Patterns Experiment," *EXARC Journal*, 3, <https://exarc.net/ark:/88735/10301>.
- Felgate, Matthew W., Simon H. Bickler, and Paul R. Murrell [2013], "Estimating Parent Population of Pottery Vessels from a Sample of Fragments: a Case Study from Inter-tidal Surface collections, Roviana Lagoon, Solomon Islands," *Journal of Archaeological Science*, 40, 2, p. 1319-1328, DOI: [10.1016/j.jas.2012.09.009](https://doi.org/10.1016/j.jas.2012.09.009).
- García Rosselló, Jaume and Manuel Calvo Trias [2019], "Etnoarqueología como experimentación: propuesta para la interpretación de los patrones de fractura," *Treballs d'Arqueologia*, 23, p. 181-202, DOI: [10.5565/rev/tda.108](https://doi.org/10.5565/rev/tda.108).
- Garrow, Duncan [2012], "Odd Deposits and Average Practice. A Critical History of the Concept of Structured Deposition," *Archaeological Dialogues*, 19, 2, p. 85-115, DOI: [10.1017/S1380203812000141](https://doi.org/10.1017/S1380203812000141).
- Gouletquer, Pierre [1979], *Préhistoire du futur. Connaître les pays est un repos*, Morlaix: Éditions Bretagnes, 112 pp.
- 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, p. 3112-3121, DOI: [10.1016/j.jas.2011.07.011](https://doi.org/10.1016/j.jas.2011.07.011).
- Malloy, Hayley [2019], *Experimental Archaeology and Formation Processes: New Experiments with Spatial Modeling*, Senior Thesis, UVM College of Arts and Sciences College, <https://scholarworks.uvm.edu/castheses/55>.
- Manen, Claire (ed.) [2022], *Le Tai (Remoulins – Gard). Premières sociétés agropastorales du Languedoc méditerranéen (6^e-3^e millénaire avant notre ère*, Toulouse: Archives d'Écologie Préhistorique, forthcoming.
- 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.
- Morin, Eugène, Arianne Boileau, and Elspeth Ready [2021], "A Refitting Experiment on Long Bone Identification," *International Journal of Osteoarchaeology*, 31, 4, p. 650-662, DOI: [10.1002/oa.2980](https://doi.org/10.1002/oa.2980).
- O'Brien, Michael J. and R. Lee Lyman [2002], "The Epistemological Nature of Archaeological Units," *Anthropological Theory*, 2, 1, p. 37-56, DOI: [10.1177/1463499602002001287](https://doi.org/10.1177/1463499602002001287).
- O'Brien, Patrick [1990], "An Experimental Study of the Effects of Salt Erosion on Pottery," *Journal of Archaeological Science*, 17, 4, p. 393-401, DOI: [10.1016/0305-4403\(90\)90004-0](https://doi.org/10.1016/0305-4403(90)90004-0).
- Orton, Clive R. [1982], "Computer Simulation Experiments to Assess the Performance of Measures of Quantity of Pottery," *World Archaeology*, 14, p. 1-20.
- Plutniak, Sébastien [2021a], *Refitting Pottery Fragments from the Liang Abu Rockshelter, Borneo*, Zenodo, DOI: [10.5281/zenodo.4719577](https://doi.org/10.5281/zenodo.4719577).
- Plutniak, Sébastien [2021b], "The Strength of Parthood Ties. Modelling Spatial Units and Fragmented Objects with the TSAR Method – Topological Study of Archaeological Refitting," *Journal of Archaeological Science*, 136, p. 105501, DOI: [10.1016/j.jas.2021.105501](https://doi.org/10.1016/j.jas.2021.105501).

- Plutniak, Sébastien [2022a], *Archeofrag: an R package for Refitting and Spatial Analysis in Archaeology*, DOI: [10.5281/zenodo.6912425](https://doi.org/10.5281/zenodo.6912425),
<https://CRAN.R-project.org/package=archeofrag>.
- Plutniak, Sébastien [2022b], “Archeofrag: an R package for Refitting and Spatial Analysis in Archaeology,” *Journal of Open Source Software*, 7, 75, p. 4335, DOI: [10.21105/joss.04335](https://doi.org/10.21105/joss.04335).
- Prignano, Luce, Ignacio Morer, and Albert Diaz-Guilera [2017], “Wiring the Past: A Network Science Perspective on the Challenge of Archeological Similarity Networks,” *Frontiers in Digital Humanities*, 4, p. 13, DOI: [10.3389/fdigh.2017.00013](https://doi.org/10.3389/fdigh.2017.00013).
- Rutkoski, Ashley Marie [2019], *An Experimental Approach to Sherd Variation*, Master of Arts Thesis, Kent State University, College of Arts and Sciences, Department of Anthropology,
http://rave.ohiolink.edu/etdc/view?acc_num=kent1555764606649139.
- Schick, Kathy Diane [1986], *Stone Age Sites in the Making. Experiments in the Formation and Transformation of Archaeological Occurrences*, British Archaeological Reports International Series, 319, Oxford: Archaeopress, 330 pp.
- Schiffer, Michael B. [1987], *Formation Processes of the Archaeological Record*, Albuquerque: University of New Mexico Press, 448 pp.
- Schiffer, Michael B. and James M. Skibo [1989], “A Provisional Theory of Ceramic Abrasion,” *American Anthropologist*, 91, 1, p. 101-115, DOI: [10.1525/aa.1989.91.1.02a00060](https://doi.org/10.1525/aa.1989.91.1.02a00060).
- Sokal, Robert R. and Charles Duncan Michener [1958], “A Statistical Method for Evaluating Systematic Relationships,” *University of Kansas Science Bulletin*, 38, p. 1409-1438.
- Villa, Paola and Jean Courtin [1983], “The Interpretation of Stratified Sites: A View from Underground,” *Journal of Archaeological Science*, 10, 3, p. 267-281, DOI: [10.1016/0305-4403\(83\)90011-0](https://doi.org/10.1016/0305-4403(83)90011-0).
- 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, New York: Academic Press, p. 539-605, DOI: [10.1016/B978-0-12-624180-8.50016-0](https://doi.org/10.1016/B978-0-12-624180-8.50016-0).