Archaeological and Anthropological Sciences (2026) 18:150
https://doi.org/10.1007/s12520-026-02510-9

RESEARCH

Reconstructing the sequentiality of adjacent flake removal scars on lithic 3D models: A curvature-based computational approach

Cheng Liu[1] · Francesco Valletta[2,3] · Javier Baena Preysler[4] · Armando Falcucci[5]

Received: 19 February 2026 / Accepted: 18 May 2026

Abstract

The traditional reconstruction of lithic reduction sequentiality relies on subjective human judgment, which is susceptible to material quality and observer experience. This study evaluates the efficacy of multiple digital geometric metrics for predicting scar sequentiality using three-dimensional scans on a technologically varied set of small-sized experimental lithic artifacts. We tested a variety of quantitative measures, including scar area size, compactness, node degree and betweenness, alongside the average and near-ridge curvature. Our results demonstrate that local curvature values measured near the common ridge provide the most reliable signature for chronological reconstruction, outperforming all other metrics. The method achieved a moderately high accuracy of up to 75%, demonstrating performance comparable to human analysts. A Generalized Linear Mixed Model (GLMM) was used to identify the factors impacting reconstruction success, showing that success is primarily determined by local scar geometry: the presence of a hinge or step was associated with higher odds of success, and opposed direction also showed a strong positive effect. In contrast, reduction method, raw material, cortex, and blank type showed limited influence on reconstruction outcomes, indicating the method's robustness to broader technological and material variation. These results validate the curvature principle as a powerful, replicable tool for analyzing large samples of small-sized artifacts, paving the way for large-scale comparative analysis of lithic assemblages.

Keywords Lithic technology · Scar sequentiality · Computational archaeology · Artifact-3D · Curvature · Automation

Introduction

This article aims to develop a quantitative method for reconstructing the sequential order of negative flake scars remaining on stone artifacts through a combination of 3D scans and novel computational toolkits. The identification of such chronological order is an integral part of the reconstruction of the chaîne opératoire (Audouze et al. 2017; Delage 2017; Porqueddu et al. 2023; Sellet 1993) and fundamental to understanding the individual technological choices (Boëda 2023) and cultural reproduction (Liu and Stout 2023) of prehistoric knappers. However, French and Anglophone approaches to lithic analysis are widely recognized as being based on different epistemological and ontological principles (Hussain 2019; Soressi and Geneste 2011). On a practical level, one such example is the "diacritical diagram", a graphical representation of stone artifacts highlighting the knapping directions of individual scars as well as their chronological relationship (Dauvois 1976: 194–201). This is the cornerstone of the French school of lithic analysis, in contrast to the attribute-based analyses more commonly

Cheng Liu and Francesco Valletta contributed equally to this work.

✉ Cheng Liu
raylc1996@outlook.com

✉ Armando Falcucci
A.Falcucci@soton.ac.uk

Francesco Valletta
fvallett@campus.haifa.ac.il

Javier Baena Preysler
javier.baena@uam.es

1 PACEA UMR 5199, CNRS & Université de Bordeaux, Pessac, France

2 Material Culture PaleoLab, School of Archaeology and Maritime Cultures, University of Haifa, Haifa, Israel

3 Department of Early Prehistory and Quaternary Ecology, Eberhard Karls University of Tübingen, Tübingen, Germany

4 Department of Prehistory and Archaeology, Universidad Autónoma of Madrid, Madrid, Spain

5 Centre for the Archaeology of Human Origins, Department of Archaeology, University of Southampton, Southampton, UK

Published online: 01 June 2026

used by Anglophone scholars (Bar-Yosef and Van Peer 2009; Shott 2007).

The creation of these diagrams relies on the reverse application of the law of superposition, which involves identifying which scars truncate others based on specific morphological features. Conventionally, this assessment is based on (1) the visual assessment of the shape of scars, where new scars should be more complete) and (2) the technological reading of scar ridge patterns aided with physical touching. The second principle assumes that older and newer scars differ in curvature where they intersect at a ridge, producing distinct tactile sensations. More specifically, the moving of one's fingers from an old scar to a new scar creates a blunt sensation, while the opposite movement a sharp one (Baena Preysler and Cuartero 2006: 150–152).

This reliance on tactile sense and the analyst's "tacit knowledge" has led many Anglophone researchers to question the replicability and objectivity of the method (Bar-Yosef and Van Peer 2009). Despite these concerns, a group of German scholars specializing in lithic analysis (Tafelmaier et al. 2022: 41; Pastoors et al. 2015: 67; Richter 1997: 192) has attempted to standardize these observations into a ranked list of five diagnostic criteria used to determine scar sequentiality (Table 1). In this study, the method is specifically restricted to the first and most reliable and quantifiable criterion: the shared scar ridge curvature pattern. Furthermore, we tested additional metrics beyond this standard list to evaluate their effectiveness in reconstruction as detailed in the method section.

Recent advancements in digital archaeology provide an opportunity to resolve the subjectivity implied in these observations by translating these analog criteria into quantifiable geometric data. 3D scanning and computational geometry now allow precise measurement of surface curvature, which can serve as a digital proxy for the analyst's manual touch. Emerging methods, such as those utilizing Multi-Scale Integral Invariants (MSII) and Integral Variant of Polylines (IIoPs), have shown promise in automating the detection of ridges and predicting the temporal relations of adjacent scars (Linsel et al. 2024, 2025). More specifically, these studies constructed a series of measurements to reconstruct the scar sequentiality, ranging from scar-level attributes such as scar area and average curvature to graph-based measurements such as degree and node betweennness, where the latter was demonstrated to display the highest accuracy (approximately 90%) after excluding all small-sized scars (e.g., taphonomic or intentional retouch). In parallel, researchers have begun to build reference collections through experimental knapping to test the reliability of results produced by human lithic analysts, including both experts and novices (Kot et al. 2025). These studies, particularly the work by Kot et al. (2025), provide crucial benchmarks by quantifying human observer error rates across different raw materials, tool types, and scar orientations, thereby establishing a reference against which automated digital methods can be evaluated.

Building upon this foundation, the present study addresses two central questions concerning the use of digital methods in lithic analysis. The first question asks whether 3D models can be used to reconstruct the relative chronology of adjacent scars using automated measurements such as near-ridge and whole-scar curvature, scar area, outline compactness, and scar node degree and betweenness, when compared with results from traditional analytical methods. This investigation explores the potential of digital modelling to quantify subtle morphological differences in scar intersections that are difficult to assess through traditional macroscopic observation. The second question examines which factors influence the success of this reconstruction, focusing on technological attributes, material properties, and morphological characteristics such as raw material type, cortex preservation, and scar orientation. To address these questions, this paper provides a quantitative and replicable assessment of a new method for establishing scar sequentiality and identifies which variables most reliably influence its performance.

Table 1 Five basic criteria for reconstructing the chronology of negative scars, which are sorted in decreasing order based on their usefulness

No.	Description
1	The younger negative cuts deeper into the raw material and exhibits a stronger concavity than the older one, especially at the common ridge.
2	At the common ridge the younger negative shows ray cracks/lancet cracks – those of the older one were overprinted when the younger one was removed.
3	The younger negative shows clearly pronounced Wallner lines in the terminal area, which are rib-shaped marks perpendicular to the crack propagation direction in brittle materials.
4	Often there are splinters on the common ridge that accompany the lancet cracks of the younger negative.
5	The outline of the younger negative follows the relief of the preceding one.

Methods and materials

The experimental assemblage

To develop reliable baseline annotations in scar chronology, we produced an experimental assemblage composed of six reduction sequences, covering two carinated cores, two preferential Levallois cores, one orthogonal-platform bladelet core, and one prismatic bladelet core (Table 2; Fig. 1). These four reduction strategies were selected to represent a

Table 2 Inventory of six nodules used for this experiment. The unit of mass is gram. Strike number refers to the number of strikes that successfully detached pieces from cores, which also corresponds to the number of debitage bags

No	Mass (gram)	Material	Technology	Strike number
1	199	Flint	Carinated	41
2	204	Basalt	Carinated	37
3	249	Basalt	Levallois (preferential)	40
4	227	Flint	Levallois (preferential)	56
5	182	Flint	Orthogonal	47
6	161	Flint	Prismatic	50

range of technological approaches associated with Middle and Upper Paleolithic assemblages, producing a variety of scar geometries, directionality patterns, and knapping intensities. This variation allows us to test whether the curvature-based method generalizes across different knapping traditions within a controlled experimental setting.

The experimental assemblage is made of two types of raw materials purchased from neolithics.com, namely flint sourced from Norfolk, UK, and fine-grained basalt. Direct percussion featuring an alternating use of hard and soft hammerstone was applied to all six core reduction sequences, where no soft organic hammers made from antler, bone or wood were involved. The hammerstones were divided into two categories based on raw material composition (Clark et al. 2020; Driscoll and García-Rojas 2014). Hard hammerstones were made from dense, smooth-grained rocks (e.g., quartzite) with relatively high mass. Soft hammerstones were made from sandy, coarser-grained rocks (e.g., sandstone) with lower mass. Because not all nodules are fully covered by natural cortex, we used silver spray paint to add a thin layer of artificial cortex on the surface of all six nodules and keep track of the portion of the nodule original surface possibly preserved on the knapped artifacts. It should also be noted that no retouching was conducted on any artifacts in our experimental assemblage.

All knapping sessions were video-recorded. During the knapping process, the knapper (author J.B.P.) paused after every single strike to allow diachronic photographic documentation of morphological changes in both cores and relevant debitage throughout the reduction sequence (Fig. 2). Additionally, we collected artifacts produced during each single successful strike in individual bags, which occasionally contain more than one detached piece. We adopted a

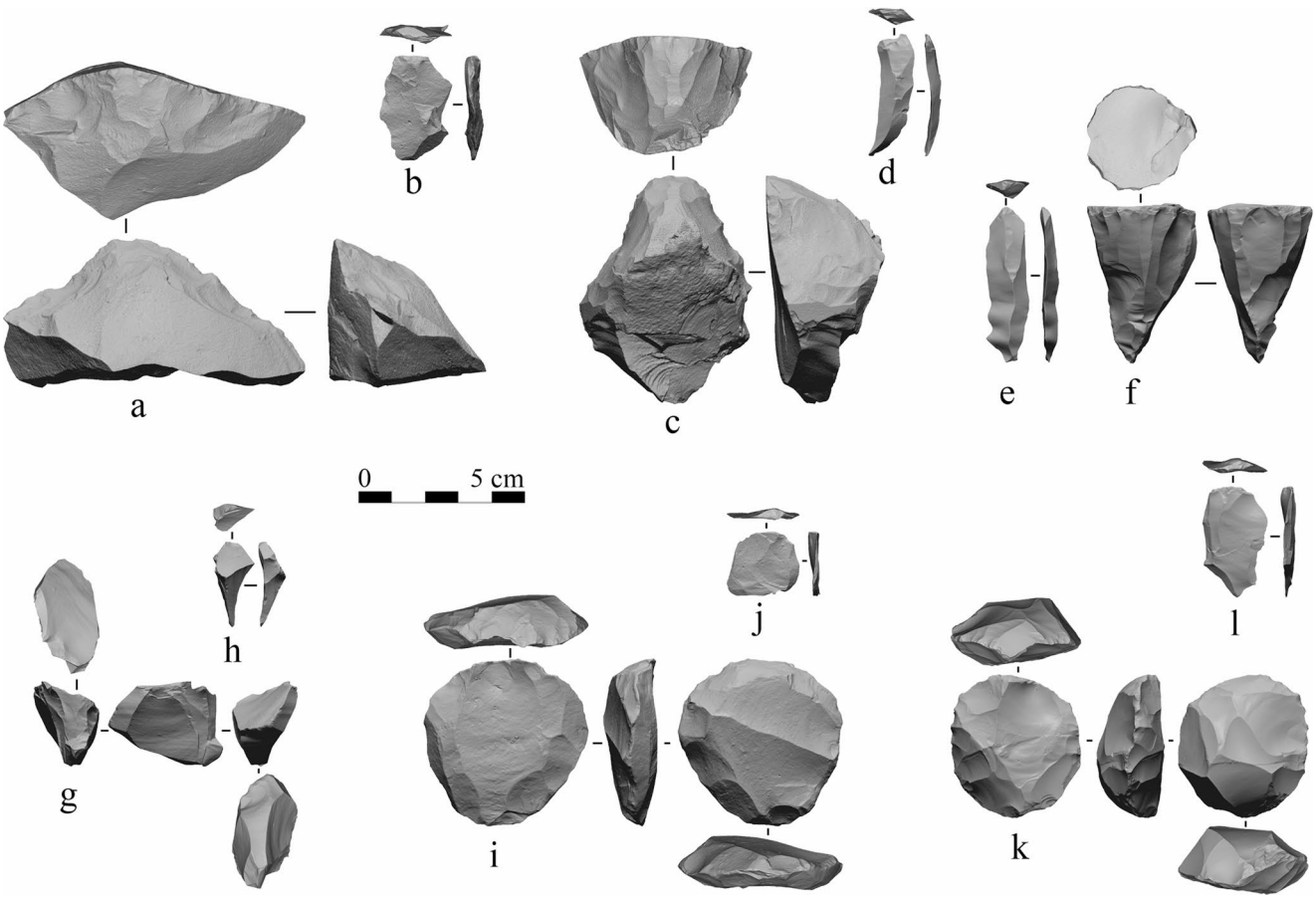

Fig. 1 3D views of reduced cores (**a**, **c**, **f**, **g**, **i**, **k**) and an example of a blank produced from each respective core (**b**, **d**, **e**, **h**, **j**, **l**). (**a**, **c**) Carinated cores; (**f**) prismatic core; (**g**) orthogonal core; (**i**, **k**) Levallois preferential cores

Fig. 2 The complete photographic record of 40 successful strikes from reduction sequence no.3, a basalt preferential Levallois core. The number on each photo represents their corresponding striking order. As an explicit example of the naming system, the final Levallois flake from the 40th strike broke into two pieces during the knapping process. These fragments are labeled respectively as 3.40_1 and 3.40_2. This system ensures that the first number identifies the reduction sequence, the second number identifies the strike order, and the final underscore distinguishes individual fragments from a single strike

labeling system to accommodate the experimental condition, where the label of a single piece of debitage is composed of two parts, namely nodule/reduction sequence number (Nn) and artifact number (An) - separated by a point (Nn.An). Specifically, the core from each nodule/reduction sequence is always labeled as Nn.0. The artifacts removed by each subsequent successful strike from the same nodule/reduction sequence Nn are labelled with progressively increasing numbers. When a strike results in the removal of multiple pieces, they are given supplementary numbers (Sn) attached to the artifact number (e.g., Nn.An_Sn).

3-D scans and ground truth annotation

Following the MicroStone protocol developed by Falcucci (2022), we adopted Artec Micro I 3D-Scanner (Artec Inc., Luxembourg) as the main platform for scanning the experimental assemblage and Artec Studio 17 as the main software for 3D scan post-processing. This results in digital models representing the artifact surface as a dense mesh of triangular facets (polygons) connecting 3D-vertices. To maximize the efficiency of extracting sequential information of scar removal, we applied the following criteria to exclude detached pieces: (1) primarily cortical; (2) too small for scanning; (3) lacking a ventral surface (i.e., angular debris). This leads to the production of 137 3D scans in total, encompassing 6 cores and 131 flakes. It should also be noted that the two carinated cores (1.0 and 2.0) were scanned by Artec Space Spider (Artec Inc., Luxembourg) given their relatively large size. These two 3D scans were later excluded from the analysis because Artifact3-D (Grosman et al. 2022) failed in the task of automatic scar segmentation on these two models with the default parameters for batch processing. We attempted to reprocess these models in Artifact3-D by adjusting the High Ridginess Factor (1→0.05) and Cost Factor (10→20) and manually merging oversegmented regions, but ultimately excluded them because the required parameter tuning falls outside a standard automated pipeline and reduces replicability and accessibility. All 3D scans went through a batch mesh simplification procedure, reducing their size to 100,000 polygons.

We used Artifact3-D (Grosman et al. 2022) v.2024 to automatically segment the 3D models in scars based on the surface curvature distribution (Richardson et al. 2013). Each scar is defined as a portion of the original 3D mesh delimited by ridges, i.e., high-curvature areas separating the scars. For each model, the algorithm produces univocal ID numbers to identify scars and ridges. The procedure is designed to over-segment the scars, that, in a second stage, should be manually merged by the user based on the close observation of corresponding specimens. Curvature-based scar segmentation has been applied in earlier work using Geomagic software (Lin et al. 2010), and a similar procedure based on the distribution of surface curvature was later developed by Linsel and colleagues (2024). Artifact3-D provides an integrated and more accessible implementation of this principle, and we opted for Richardson's approach as it was already built into the software. Subsequently, based on the

knapping video, photos after every single successful strike, and photo-assisted refitting of debitage, we established the ground truth annotation of scar chronology referring to the automatically assigned ID numbers.

Automatic definition of scar sequentiality

Throughout this paper, a scar dyad refers to any pair of adjacent scars that share a common ridge. For each artifact, the set of scar dyads comprises all pairwise combinations of neighboring scars, and the analysis attempts to determine the chronological order within each dyad independently based on the following scar-related parameters extracted from automatically segmented 3D models of the artifacts.

Whole-scar curvature: The basic assumption for using this feature is the initial part of the first criterion presented in Table 1: i.e., that younger scars cut more deeply into the raw material and exhibit a stronger concavity. At each point of a 3D surface, the degree of concavity/convexity can be calculated based on the main curvature value, i.e. the larger of the two principal curvature values computed from the two perpendicular tangent planes at each surface vertex. We calculated the main surface curvature value of each vertex of the 3D mesh with the procedure included in Artifact3-D (Cohen-Steiner and Morvan 2003; Grosman et al. 2022); positive values denote points in which the main curvature is convex, negative values concave. Consequently, the scar with the lower average main curvature has a more convex surface and is, therefore, defined as the younger one.

Near-ridge curvature: According to the first diagnostic criterion presented in Table 1, near the ridge at the intersection between two scars, the later (more complete) one has a higher concavity value than the earlier (less complete) one. In our procedure, the shared ridge between the two scars is defined as the ridge (or set of ridges) included in the border of both scars. For each of the two scars, the average main curvature value is then calculated based on the main curvature of the vertices within a set distance (d) from the shared ridge. The scar with the lower value is expected to preserve more of the convex surface near its original border and is therefore defined as the more recent.

Scar area: The rationale underpinning this criterion is that the more recent scar partially overlaps with the earlier one, reducing its surface. We calculate the scar area with the procedure integrated within Artifact3-D (Grosman et al. 2022). More extensive scars are, therefore, expected to be younger.

Scar outline compactness: The outline of older scars is expected to become more complex (less compact) as they are intercepted by younger scars that create "dents" in their original perimeter. The compactness of a closed polygon (e.g., a scar outline) can be measured as the ratio between the circumference of a circle with the same area and the actual length of its perimeter. We measured the compactness of the outline of adjacent scars based on the procedure integrated within Artifact3-D (Grosman et al. 2022). Younger scars are expected to have a more compact outline than older ones.

Scar node degree and betweenness centrality: These features were included since they were particularly effective in correctly predicting the scar order in the experiment conducted by Linsel et al. (2025). They are based on a graph, where each scar is represented as a node connected by undirected edges to its neighbors. The degree of each scar equals the number of nodes directly connected to it, i.e., its neighbors. The scar betweenness centrality equals the frequency with which a node/scar is included in the shortest path (in terms of number of edges) between two other nodes/scars.

Lithic attributes

To understand how the accuracy of the curvature-based scar chronology reconstruction method is impacted by different factors, we recorded a series of lithic attributes as listed in Table 3. One factor is the orientation pattern of adjacent scars (Lin et al. 2024), for which we developed a protocol of documentation (Fig. 3) and conducted an inter-observer replicability (Pargeter et al. 2023; Timbrell et al. 2022) test among three of the authors (C.L., F.V., and A.F.). In the replicability test, we use an online tool (https://www.gigacalculator.com/randomizers/random-picker.php) to randomly select 5 specimens from the whole list. The results of the inter-rater agreement test indicate a relatively high Fleiss's

Table 3 Definition of recorded lithic attributes

Attribute	Definition
Method	The overall reduction sequence utilized. Levels include Carinated, Levallois (preferential), Orthogonal, and Prismatic.
Material	The raw material of the studied piece. Levels are basalt and flint.
Cortex	Describes the preservation of cortex on the corresponding piece. Levels are binary (preserved or not preserved).
Blank	The basic technological and morphological classification of the blank. Levels include core, flake, blade (i.e., a flake with a length at least twice its maximum width; Inizan et al. 1995), and bladelet (i.e., a blade with a maximum width equal or below 12 mm; Tixier 1963).
Hinge/Step	Records the existence of a hinge or step termination (Cotterell and Kamminga 1987) feature present on the corresponding adjacent scar ridge. Levels are binary.
Orientation	Defines the knapping orientation of the corresponding scar dyad. Levels are A, B, and C (see Fig. 3).

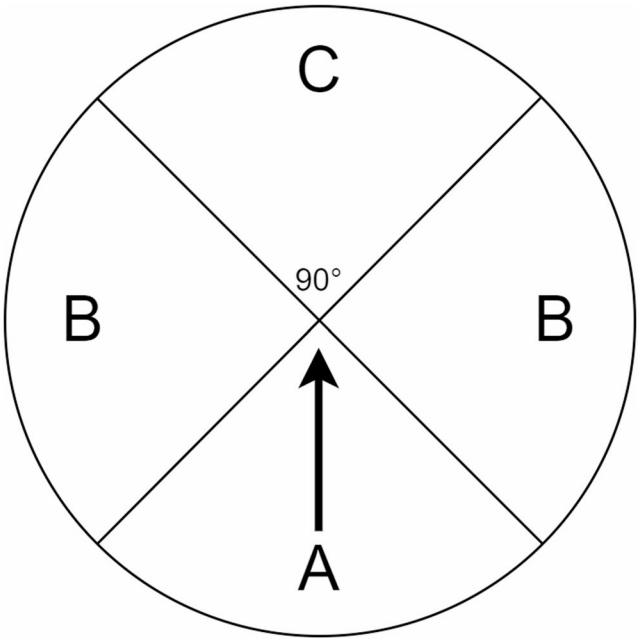

Fig. 3 Graphic illustration of the three categories of knapping orientation used to classify scar dyads. Orientation **A**: both scars in the dyad were detached from the same directional quadrant (angular difference approximately 0–45 degrees). Orientation **B**: scars were detached from adjacent directional quadrants (approximately 45–135 degrees). Orientation **C**: scars were detached from opposite directions (approximately 135–180 degrees). The reference direction (as indicated by the arrowhead) is defined as the primary knapping direction of the most recent removal on the artifact as determined from the ground truth annotation

kappa value (0.68) and an extremely low p-value (<0.001), representing modest agreement according to Fleiss classification (Fleiss et al. 2003: 598–626).

Statistical analysis

The primary analysis of the correlation between the success rate of the automatic sequentiality definition and the recorded technological features was conducted using a Generalized Linear Mixed Model (GLMM) with a binomial family and a logit link function. This multi-level approach was required to address the dependency among observations, as multiple scar dyad measurements were obtained from the same debitage piece and shared the same levels for the Method, Material, Cortex, and Blank attributes. The model included a random intercept grouped by the unique debitage ID to allow the baseline probability of chronological success to vary randomly across individual pieces. All models were fitted using the *lme4* R package v.1.1–37 (Bates et al. 2015) with the bobyqa optimization algorithm to ensure robust and reliable convergence.

We initially constructed a full model to explore potential interaction effects between the technological attributes and the buffering zone distance (d). In R notation, this structure is: success ~ (Method + Material + Cortex + Blank_type + hinge_step + orientation) * Distance + (1 | ID), where success is the binary outcome (1 = correct, 0 = incorrect) for each scar dyad. Following the initial exploration of the full model, we addressed concerns regarding potential overfitting and the inclusion of uninformative predictors by performing an automated model selection process. This was executed using the dredge function from the *MuMIn* package v.1.48.19 (Bartoń 2026), which identifies the most parsimonious model based on the Akaike Information Criterion corrected for small sample sizes (AICc). This selection process prioritized models that balanced explanatory power with structural simplicity.

To further understand the influence of local geometric features on the accuracy of the method, we analyzed the absolute difference in near-ridge curvature between two adjacent scars. A Wilcoxon rank-sum test was employed to compare curvature values between correctly and incorrectly classified dyads across three distance parameters. This non-parametric approach allowed us to determine if the degree of geometric distinctness between adjacent scars significantly affected the success of the chronological reconstruction.

Following the principle of research reproducibility and data transparency, the code and data for the statistical analysis are fully open access at the author's GitHub repository (https://github.com/Raylc/scar-stratigraphy). The developed procedure for reconstructing the sequentiality of scar dyads segmented with Artifact3-D is also available as a MATLAB code to be compiled in the same GitHub repository. For the purpose of creating a reproducible computational environment, we used the *renv* package v.1.0.7 (Ushey and Wickham 2024), which snapshotted the state of the project library in the lockfile. All 3D models and subsequent data obtained in Artifact3-D are published in an open-access Zenodo repository (https://doi.org/10.5281/zenodo.18261894).

Results

The results section presents a quantitative evaluation of the digital method for reconstructing scar chronology. First, we provide an assessment of the accuracy of individual metrics to determine which automated measurements most reliably distinguish the age of adjacent scars. This initial analysis addresses the first research question by comparing digital attributes, such as near-ridge curvature, scar area, and node degree, against the known ground truth of the reduction sequence. Following this, we use a Generalized Linear Mixed Model (GLMM) to address the second research question, identifying how specific technological and morphological factors, including orientation and the presence of hinge or step fractures, influence the success rate of the

reconstruction. To further understand the influence of local geometric features on the accuracy of the method, we also examine how the degree of geometric contrast between adjacent scar borders relates to successful chronological identification. By analyzing these variables, we aim to define the conditions under which digital scar chronology reconstruction is most effective and replicable.

General accuracy evaluation

The accuracy values reported in this section reflect the performance of each metric when used alone as the decision rule for a pairwise scar dyad comparison: the metric assigns the younger label to the scar with the higher curvature (or larger area, more compact outline, etc.), and accuracy is the proportion of dyads for which this assignment matches the ground truth. These single-metric accuracy values are distinct from the GLMM analysis, which examines which attributes predict success across all dyads simultaneously.

The near-ridge curvature reaches the highest accuracy (74.2% on average), while other measurements tested here display a significantly lower success rate, often near 50%. These include scar average curvature (60.8%), scar area size (47.5%), scar outline compactness (55.3%), scar node degree (43.5%), and scar node betweenness (50.8%). For the next step, we will focus our analysis of the near-ridge curvature as it displays the highest accuracy, which is also deeply rooted in the first criterion of technological reading that "the younger negative cuts deeper into the raw material and shows a stronger concavity than the older one (Tafelmaier et al. 2022: 41)." More specifically, we measured the average curvature in a buffering zone extended from the ridge between two adjacent scars, for which we set three different distance (d) parameters: 0.5 mm, 1 mm, and 2 mm (Fig. 4). It was found that the 1 mm buffering zone has the highest accuracy (75%) in reconstructing the chronology of flake removals, while the 0.5 mm (73.9%) and 2 mm (73.7%) buffering zones feature slightly lower performance scores.

Factors affecting accuracy

The final dataset is composed of 529 scar dyads drawn from 106 artifacts, including 4 cores, 73 flakes, 13 blades, and 16 bladelets. The discrepancy between the total number of artifacts analyzed and the available scans stems from the

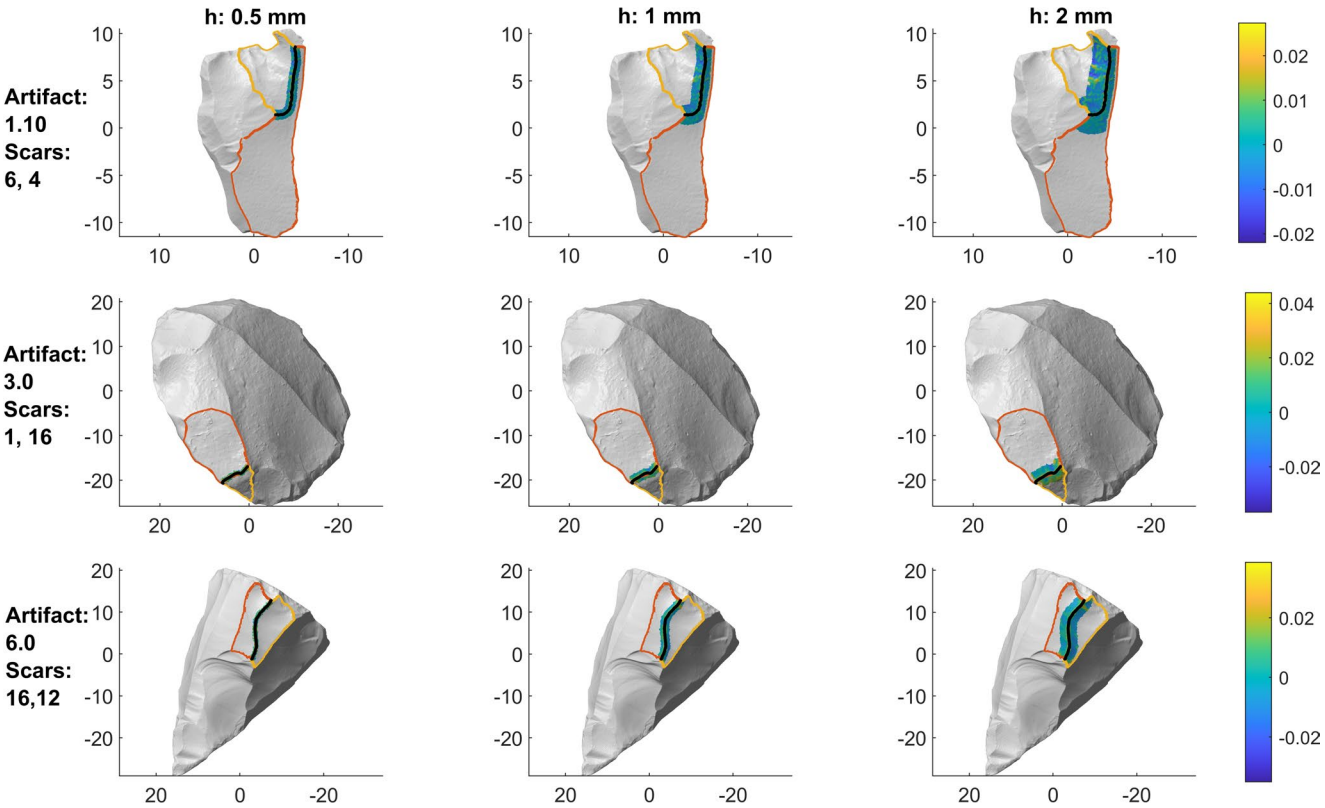

Fig. 4 Three examples of scar ridge curvature measurements with different buffering zone distance (d) parameters. 1a) Flake 1.10, ridge curvature between scar 4 and 6, 0.5 mm; 1b) Flake 1.10, ridge curvature between scar 4 and 6, 1 mm; 1c) Flake 1.10, ridge curvature between scar 4 and 6, 2 mm; 2a) Core 3.0, ridge curvature between scar 1 and 16, 0.5 mm; 2b) Core 3.0, ridge curvature between scar 1 and 16, 1 mm; 2c) Core 3.0, ridge curvature between scar 1 and 16, 2 mm; 3a) Core 6.0, ridge curvature between scar 12 and 16, 0.5 mm; 3b) Core 6.0, ridge curvature between scar 12 and 16, 1 mm; 3c) Core 6.0, ridge curvature between scar 12 and 16, 2 mm

Table 4 GLMM results with odds ratios. Full model results are reported in Supplementary Table S1

Term	Coefficient Log Odds	Standard Error	Z Value	*P*Value	Odds Ratio	Lower 95 CI	Upper 95 CI
(Intercept)	1.813	0.737	2.460	0.014	6.128	1.446	25.971
Blank_typebladelet	−1.256	0.724	−1.734	0.083	0.285	0.069	1.178
hinge_step	0.771	0.353	2.184	0.029	2.163	1.082	4.322
orientationC	1.295	0.569	2.275	0.023	3.652	1.197	11.144
Distance2mm	−1.485	0.866	−1.715	0.086	0.227	0.042	1.236
MethodLevallois (preferential): Distance2mm	1.226	0.559	2.195	0.028	3.409	1.140	10.194
MethodOrthogonal: Distance2mm	1.629	0.705	2.310	0.021	5.098	1.280	20.303

fact that dyads for which any attribute value could not be determined (mainly orientation) were excluded from the analysis. The Multi-Level Logistic Regression analysis identified two attributes with meaningful effects on the success of adjacent scar chronology reconstruction (Table 4; Fig. 5), namely hinge/step and orientation. The estimated variance of the random intercept for the debitage piece ID was 0.776 (s.d. = 0.881), confirming that the baseline success rate varies substantially across the 106 individual debitage pieces after removing observations containing missing value. This further validates the use of the multi-level modeling approach.

The attribute hinge_step showed a positive effect on reconstruction success ($\beta = 0.771$, $p \approx 0.029$). The odds of successful reconstruction were approximately 116.3% higher (Odds Ratio ≈ 2.163) when a hinge or step was present on the adjacent scar ridge compared to when it was absent. The knapping orientation of the corresponding scar dyad also influenced reconstruction success. Compared to the reference Orientation A (the two adjacent scars

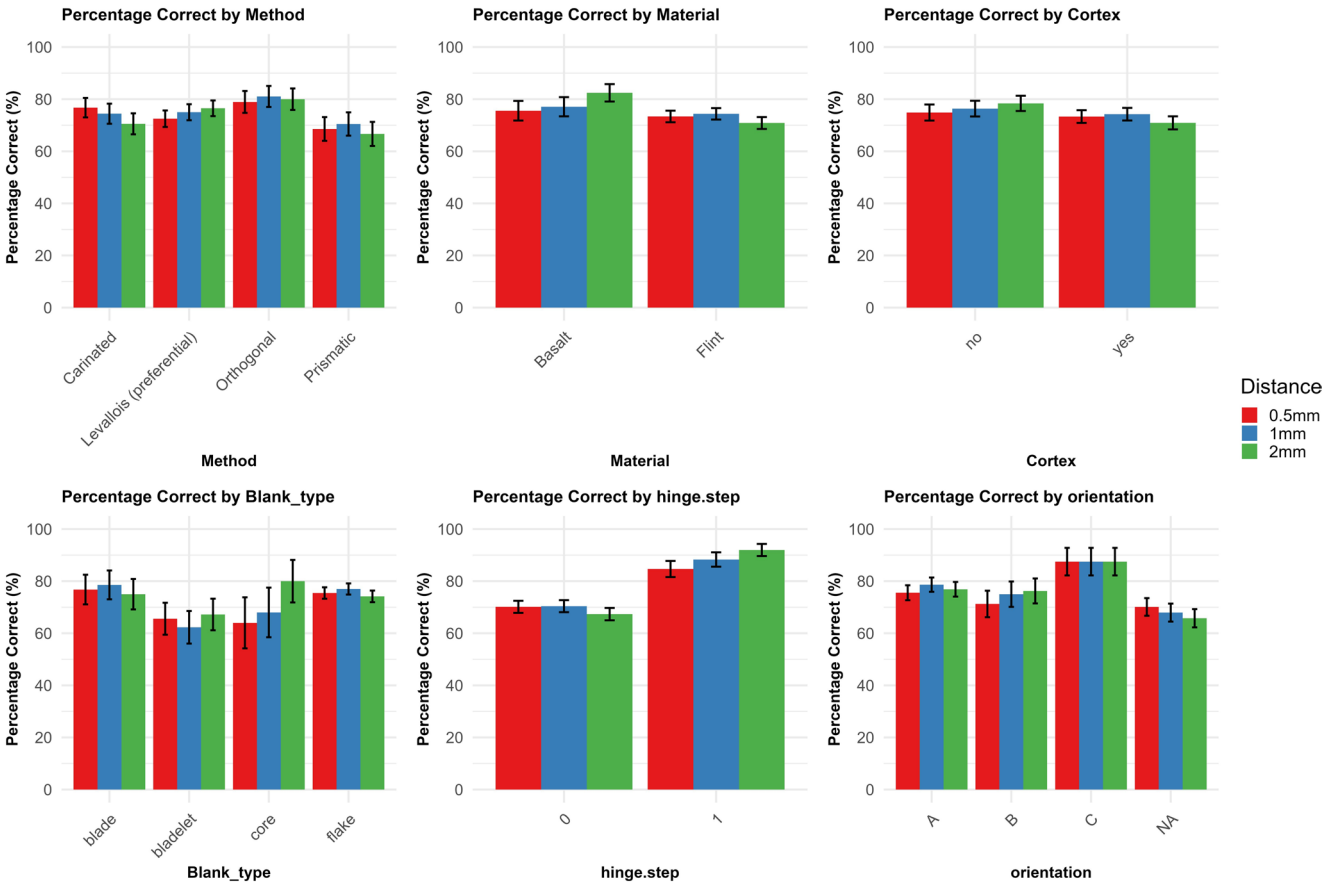

Fig. 5 Bar plot displaying mean accuracy rates at three buffering zone distance parameters (0.5 mm, 1 mm, 2 mm) under different attribute conditions. Each panel shows results for one attribute, with bars representing individual levels of that attribute. Error bars represent the Standard Error of the Mean (SEM). Accuracy is calculated as the proportion of scar dyads for which the near-ridge curvature correctly identified the younger scar

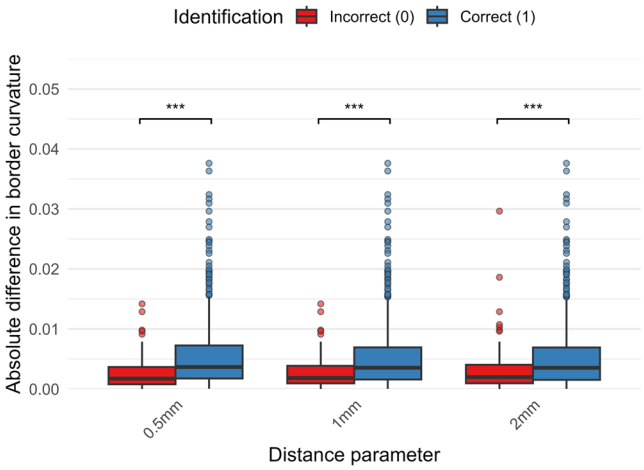

Fig. 6 Box plot displaying the mean absolute difference in near-ridge curvature at three distance parameters in correct (1) and incorrect (0) chronological reconstructions. Brackets with three asterisks indicate differences within each group supported at $p<0.001$

are detached from the same 90 degree direction range), Orientation C ($\beta = 1.295$, $p \approx 0.0229$) was associated with higher odds of success. The odds of successful reconstruction were approximately 265.2% higher (Odds Ratio ≈ 3.652) when the two adjacent scars are detached from opposite directions.

The analysis further indicated that while the buffering zone distance did not influence the effects of most technological and morphological factors ($p > 0.05$), it did have a specific impact on certain reduction methods. The model identified interactions between the reduction method and the 2 mm buffering zone. Both Levallois preferential reduction ($\beta = 1.226$, $p \approx 0.028$) and orthogonal reduction ($\beta = 1.629$, $p \approx 0.021$) demonstrated higher accuracy at the 2 mm distance compared to carinated reduction. In contrast, the 2 mm distance parameter itself ($\beta = -1.485$, $p \approx 0.086$) and the bladelet blank type ($\beta = -1.256$, $p \approx 0.083$) provided weaker evidence of an effect, suggesting a potential decrease in reconstruction success under those specific conditions. The estimated effects of Material, Cortex, and Orientation B remained small, showing limited influence on the success of the reconstruction in this model.

The visualization of the mean accuracy by attribute levels (Fig. 5) supports the effect patterns identified by the Generalized Linear Mixed Model (GLMM). The rigorous finding that the existence of a hinge/step significantly increases success (Odds Ratio ≈ 2.163) is mirrored in the graph by a visibly higher mean accuracy for the corresponding non-reference level of that attribute. Similarly, the statistical finding that Orientation C is positively associated with success (Odds Ratio ≈ 3.652) is visually supported by the mean accuracy bar for Orientation C being distinctly higher than the reference Orientation A.

Furthermore, we examined the absolute difference in curvature between the two scar borders across all three distance parameters. This analysis aimed to determine if the degree of geometric contrast between adjacent scars significantly influences the likelihood of a successful chronological reconstruction. The results of a Wilcoxon rank-sum test (Mann-Whitney U test) compare the absolute difference in near-ridge curvature values between scar pairs in correctly and incorrectly classified dyads, which are two independent groups. Each point in the box plot represents one scar dyad (Fig. 6). The difference in near-ridge curvature is substantial across all measured scales. At the 0.5 mm distance, correctly classified dyads showed significantly higher curvature contrast than incorrectly classified dyads ($W = 16760$, $n_{incorrect} = 138$, $n_{correct} = 391$, $p > 0.0001$). This pattern remained consistent at the 1 mm distance ($W = 17917$, $n_{incorrect} = 132$, $n_{correct} = 397$, $p < 0.0001$) and the 2 mm distance ($W = 19663$, $n_{incorrect} = 139$, $n_{correct} = 390$, $p < 0.0001$). Across all three parameters, the absolute difference in curvature for correct identifications (0.0055–0.0057) was consistently higher than for incorrect identifications (0.0025–0.0030).

Discussion

The assessment of the novel automated curvature-based method introduced here demonstrates a high success rate in reconstructing the relative chronology of flake removals, achieving a peak accuracy of 75% in the 1 mm buffering zone. When benchmarked against human observers, this performance is highly competitive. Kot et al. (2025) reported that the average error rate across all human participants was 21%, corresponding to an average accuracy of 79%. Even the lithic specialists in that study, who are characterized by lower error rates, performed with an average accuracy of 80%, with only the most experienced achieving 85% accuracy. This comparison suggests that the automated method already performs at a level comparable to archaeology students trained in technological reading, who show an error rate of about 25%, and in some cases approaches the performance of trained specialists, despite the fact that the debitage analyzed here is generally smaller and potentially more difficult to interpret than the bifacial artifacts that often form the focus of traditional scar chronology studies. The consistency of the method, which is not affected by observer experience or fatigue, thus presents a substantial advantage for large-scale lithic studies.

Although the curvature principle is a strong predictor of chronological success, it is not without limits, and its failure cases are important for understanding the boundaries

of the method. Failures of the algorithm often occur when two adjacent scars intersect at a location where the ridge between them is highly flattened. In these cases, the geometric differences in curvature between the older, flatter scar and the younger, more convex scar are minimal and fall below the detection threshold of the algorithm. This information loss can be exacerbated by technological behaviors, such as extensive platform preparation or subsequent edge trimming, which mechanically alter the original intersection geometry. In addition, the accuracy of the final output may be limited by low-quality scanning data, the reliability of interpolation methods used to fill surface gaps, or limitations of the computational algorithms used for curvature calculation. Because the semi-automated method relies on localized ridge geometry, any factor that blurs or smooths the transition between scars reduces the reliability of curvature measurements and weakens the resulting chronological inference.

Our multilevel modeling results show that specific dyad-level attributes, namely hinge or step termination and scar orientation, are more influential than artifact-level attributes in predicting reconstruction success, indicating that the method is sensitive to the physical geometry produced during knapping. The finding that hinge or step terminations significantly increase the odds of successful reconstruction supports a model in which technical accidents, rather than clean removals, generate more distinct chronological indicators. A hinge or step termination results from an abrupt fracture that leaves a sharply defined, vertical ridge on the debitage piece. This pronounced morphology creates strong relief and high contrast in surface curvature at the point where the scars intersect. The resulting increase in curvature contrast between the two scars makes the removal order less ambiguous and significantly improves reconstruction accuracy.

The strong positive effect of Orientation C, representing opposed scar directions, is notable, as this configuration increases the odds of success by 265.2%, despite traditional analyses sometimes linking opposed scars to higher human error rates (Kot et al. 2025 : 8–9). This result suggests that the curvature-based method detects a specific and reliable geometric signature in opposed removals. While opposed scars create a complex intersection pattern, the directional conflict may force the scars to meet at a highly consistent and well-defined angle of intersection that the automated algorithm, unlike a human observer, is exceptionally adept at quantifying and resolving. Therefore, for the curvature analysis method, the geometry of opposed scars provides more definitive chronological data than other orientations.

In addition to the primary effects, the full model identified four additional predictors that provide further insight into the operational limits of the curvature-based method. The results revealed evidence for two specific interaction terms: Levallois preferential reduction and orthogonal reduction both demonstrated higher reconstruction accuracy when analyzed with a 2 mm buffering zone. This suggests that for these specific reduction strategies, a wider measurement area may better capture the diagnostic concavity of the scar ridges compared to narrower parameters. Furthermore, the model indicated two marginal effects involving the 2 mm distance parameter and the bladelet blank type. The negative trend for bladelets suggests that their reduced surface area and specific geometry may approach the limits of the current algorithmic resolution. However, it is important to note that with the exception of the bladelet category, these variables were excluded from the final parsimonious model during the AICc selection process. This indicates that while these factors are informative for specific technological contexts, they possess less global predictive weight than the presence of hinge fractures or scar orientation. Further research is necessary to fully explore the underlying reasons for these conditional relationships.

The moderately high performance achieved by the semi-automated method highlights several points of divergence from human-based studies, particularly regarding the influence of reduction method and raw material. Kot et al. (2025) found that raw material quality was a highly significant predictor of human error, with uniform, highly crystalline flints being easier to analyze than coarse, heterogeneous materials. They also observed significant differences in error rates related to tool type, with large bifacial tools producing fewer errors than discoidal cores. In contrast, we found that the estimated effects of reduction method and raw material on the success of the semi-automated method were small and showed limited explanatory value. This could potentially indicate that the digital analysis, by focusing strictly on the geometry of the scar intersection, is less sensitive to the technological and raw material variables that often challenges human visual judgment. It should be noted, however, that the absence of a significant effect for these variables may partly reflect limited statistical power due to the small sample size in our controlled experimental design. We therefore interpret these results as showing no detectable effect under current experimental conditions, rather than as evidence that reduction method and raw material are categorically irrelevant to the curvature-based method. Moreover, a more direct validation would compare the automated method and human observers on the same dataset, identifying whether failures occur on the same dyads, which should be a priority for future work.

Beyond the comparisons with human performance, it is necessary to consider how these results align with other digital approaches, such as the metrics reported by Linsel et al. (2025). Their study utilized a combination of graph-based

data and surface measurements to predict scar order, yet the discrepancy between our findings and their results likely stems from fundamental differences in the establishment of ground truth and the generalizability of the metrics employed. A critical limitation of the Linsel dataset is its reliance on a lithic analyst's technological reading as the primary ground truth. Even in so-called self-created experimental sequences, the inclusion of archaeological artifacts, for which the true removal sequence cannot be known with certainty, introduces uncertainty that may reduce the reliability of the training data. In contrast, our study uses a strictly controlled experimental sequence in which the removal order is documented with certainty, providing a stronger baseline for validating digital proxies.

Furthermore, several high-performing measurements in the Linsel study lack a firm foundation in the mechanical principles of lithic reduction, making them susceptible to bias and limited in their generalizability across different technologies. For instance, graph-based metrics such as node degree are highly dependent on the specific reduction method. In a preferential Levallois context, the final flake removal creates a central scar that intersects with nearly all previous preparatory scars; thus, a high degree naturally correlates with a later chronological position. However, in prismatic core technologies used for blade and bladelet production, most scars are only connected to two adjacent removals in a linear sequence. In this context, node degree provides little chronological information. Other measures, such as average scar curvature, are frequently disrupted by adjacent scars during dyad-wise comparisons, while area size remains unreliable because a late-stage "clearing" flake can easily be larger in absolute size than the earlier scars it removes to rejuvenate the debitage surface. Finally, compactness, a metric inherent to Artifact-3D but not used in Linsel et al. (2025), fail because they measure how closely a scar approximates a perfect circle rather than directly capturing the completeness or convexity that serves as the true proxy for chronological order.

The present analysis holds particular relevance for the study of Middle Paleolithic and Upper Paleolithic technologies. Unlike traditional scar chronological studies that often rely on larger or bifacial artifacts (Baena Preysler et al. 2018; Frick and Floss 2017; Shipton et al. 2025), our pilot project focuses on the small-sized reduction sequences, especially those associated with laminar technologies (blades, bladelets, and associated small debitage flakes). These assemblages are often characterized by a significantly larger sample size in archaeological contexts (Falcucci et al. 2022; Lombao et al. 2023; Valletta et al. 2020), sometimes beyond the capacity of individual human lithic specialists of conducting a detailed diacritical analysis. Given that no detectable effect of reduction method or raw material was found under the current experimental conditions, its application to these smaller, more numerous artifacts offers a powerful new tool, especially when combined with new batch-scanning protocols that enable the creation of large numbers of 3D models within a single micro-CT scanning session (Falcucci et al. 2022; Göldner et al. 2022).

Despite the strong performance of the semi-automated approach under controlled conditions, the efficacy of chronological reconstruction remains constrained by the reality of the archaeological record (Vaesen and Hussain 2026). This limitation is highlighted by previous research demonstrating that only 39.9% of detached removals are preserved on the surface of cores and bifaces, and only 14.1% of the chronological relationships between sequences of removals can be identified (Kot et al. 2024). This inherent information loss necessitates a better methodological response. Our current methodology represents a foundational step, but its reconstruction is limited to dyad-based or pairwise chronological relationships. A key future direction is the development of algorithms that can connect the chronological relationships between different dyads to eventually generate a complete Harris matrix representing all available scars on a single debitage piece (Tafelmaier et al. 2022: 44–46). Furthermore, to fully enhance the method's reliability, it must integrate additional geometric criteria that human experts utilize. This includes expanding the analysis beyond curvature alone to assess overall scar morphology, such as convexity as a measurable indicator of scar completeness in areas adjacent to the ridge. These methodological expansions are necessary to fully capture the holistic geometric data required for comprehensive technological interpretation. At last, the combination of knapping video recording and high-resolution three-dimensional scans presents a unique opportunity to integrate different data modalities in era of artificial intelligence (Miao et al. 2025), namely knapping process (videos) and knapping product (3d scans), to make deeper inferences between visible material culture and invisible human behaviors, ultimately, mitigate the challenges posed by limited preserved information (Liu 2024).

Several limitations of this study should be acknowledged. First, all reduction sequences were produced by a single knapper, which means that individual knapping style cannot be disentangled from the effects of technology or raw material. Second, the experimental assemblage covers only two fine-grained raw materials; the performance of the method on coarser or more heterogeneous materials remains to be tested. Third, the sample of six reduction sequences provides a limited basis for generalizing across the full range of technological variability in the Paleolithic record. We therefore present these findings as a validation under controlled conditions and encourage future work to test the method on larger and more varied assemblages, including archaeological

material where the distribution of scar types may differ substantially from that in experimental collections.

Conclusion

Our findings address the core research questions of this study by confirming the efficacy of the semi-automated method and isolating the dominant predictive factors. First, the question of whether three-dimensional scans can be used to reconstruct the relative chronology of adjacent scars is answered affirmatively. Among the multiple metrics evaluated, including scar area size, compactness, node degree and betweenness, local curvature values measured near the common ridge proved to be the most robust predictive factor. The semi-automated method achieved a peak accuracy of 75% in the 1 mm buffering zone. This performance level is comparable to the average accuracy (79%) of human lithic analysts performing the same task, demonstrating that the proposed procedure offers a reliable and non-subjective means of chronological reconstruction. Second, the question of what factors impact the success rate of such reconstruction is answered by the statistical significance of two key scar-level attributes, suggesting that the success of the semi-automated reconstruction is primarily predicted by the geometry created during the removal event. The limited influence of broader technological and material variables indicates that the method is primarily sensitive to local scar geometry rather than global typological attributes. Specifically, the method's success is shown to be largely unaffected by the reduction method adopted or the quality of the raw material, factors which typically increase error rates for human observers. We note, however, that both raw materials tested here are fine-grained, and the performance of the method on coarser materials remains an open question. Future work should focus on integrating additional geometric criteria, such as scar completeness and outline convexity, to fully match the holistic decision-making capacity of human experts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12520-026-02510-9.

Acknowledgements This project is generously funded by the Global Fellowship of Senckenberg Society for Nature Research (C.L.), the Teach@Tübingen Fellowship (F.V.), the National Project reference PID2022-138590NB-C42 (J.B.P.), and the Global Marie Skłodowska-Curie grant No. 101152531 (A.F.; https://doi.org/10.3030/101152531). We would like to thank Nicholas Conard for providing access to the 3D scanning devices and for hosting C. L.'s Senckenberg project and F. V.'s Teach@Tübingen Fellowship.

Authors' contributions Cheng Liu : Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Funding acquisition, Writing – original draft, Writing – review & editing. Francesco Valletta : Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. Javier Baena Preysler : Resources, Writing – review & editing. Armando Falcucci : Data curation, Investigation, Visualization, Resources, Writing – review & editing.

Data availability All 3D models used in this study and subsequent data obtained in Artifact3-D are published in an open-access Zenodo repository (https://doi.org/10.5281/zenodo.18261894). The R code and data for the statistical analysis alongside the developed MATLAB procedure for reconstructing the sequentiality of segmented scar dyads are fully open access at the author's GitHub repository (https://github.com/Raylc/scar-stratigraphy).

Declarations

Generative AI and AI-assisted technologies in the manuscript preparation process During the preparation of this work the author(s) used ChatGPT and Grok in order to polish the language and assist in the writing of the R script. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Competing interests The authors declare no competing interests.

References

Audouze F, Bodu P, Karlin C, Julien M, Pelegrin J, Perlès C (2017) Leroi-gourhan and the chaîne opératoire: A response to delage. World Archaeol 49(5):718723. https://doi.org/10.1080/00438243.2017.1416012

Baena Preysler J, Cuartero F (2006) Más allá de la tipología lítica: Lectura diacrítica y experimentación como claves para la reconstrucción del proceso tecnológico. Zona Arqueológica 7:145–160. https://dialnet.unirioja.es/servlet/articulo?codigo=2259131

Baena Preysler J, Torres Navas C, Sharon G (2018) Life history of a large flake biface. Q Sci Rev 190:123–136. https://doi.org/10.1016/j.quascirev.2018.04.015

Bar-Yosef O, Van Peer P (2009) The chaîne opératoire approach in middle paleolithic archaeology. Curr Anthropol 50(1):103–131. https://doi.org/10.1086/592234

Bartoń K (2026) MuMIn: multi-model inference. R package version 1.48.19

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 67:1–48. https://doi.org/10.18637/jss.v067.i01

Boëda É (2023) Techno-logic & technology: A paleo-history of knapped lithic objects. Routledge. https://doi.org/10.4324/9781003359081

Clark JE, Pastrana A, Woods JC, Patten B (2020) Hammerstones Arqueología 60:1748. https://revistas.inah.gob.mx/index.php/arqueologia/article/view/17027

Cohen-Steiner D, Morvan J-M (2003) Restricted delaunay triangulations and normal cycle. SoCG03: Annual ACM Symposium on Computational Geometry. 312–321. https://doi.org/10.1145/777792.777839

Cotterell B, Kamminga J (1987) The Formation of Flakes. Am Antiq 52(4):675–708. https://doi.org/10.2307/281378

Dauvois M (1976) Précis de dessin dynamique et structural des industries lithiques préhistoriques. Pierre Fanlac

Delage C (2017) Once upon a time… the (hi)story of the concept of the *chaîne opératoire* in French prehistory. World Archaeol 49(2):158–173. https://doi.org/10.1080/00438243.2017.1300104

Driscoll K, García-Rojas M (2014) Their lips are sealed: Identifying hard stone, soft stone, and antler hammer direct percussion in palaeolithic prismatic blade production. J Archaeol Sci 47:134141. https://www.sciencedirect.com/science/article/pii/S0305440314001435

Falcucci A (2022) MicroStone: Exploring the capabilities of the Artec Micro in scanning stone tools. https://doi.org/10.17504/protocols.io.81wgb6781lpk/v1

Falcucci A, Karakostis FA, Göldner D, Peresani M (2022) Bringing shape into focus: Assessing differences between blades and bladelets and their technological significance in 3D form. J Archaeol Science: Rep 43:103490. https://doi.org/10.1016/j.jasrep.2022.103490

Fleiss JL, Levin B, Paik MC, Fleiss J (2003) *Statistical methods for rates & proportions*, 3rd edition. Wiley-Interscience

Frick JA, Floss H (2017) Analysis of bifacial elements from grotte de la verpillière i and II (germolles, france). Quatern Int 428:3–25. https://doi.org/10.1016/j.quaint.2015.10.090

Göldner D, Karakostis FA, Falcucci A (2022) Practical and technical aspects for the 3D scanning of lithic artefacts using microcomputed tomography techniques and laser light scanners for subsequent geometric morphometric analysis. Introducing the StyroStone protocol. PLoS ONE 17(4):e0267163. https://doi.org/10.1371/journal.pone.0267163

Grosman L, Muller A, Dag I, Goldgeier H, Harush O, Herzlinger G, Nebenhaus K, Valetta F, Yashuv T, Dick N (2022) Artifact3-D: New software for accurate, objective and efficient 3D analysis and documentation of archaeological artifacts. PLoS ONE 17(6):e0268401. https://doi.org/10.1371/journal.pone.0268401

Hussain ST (2019) The French-Anglophone divide in lithic research: A plea for pluralism in Palaeolithic Archaeology [PhD thesis]. https://hdl.handle.net/1887/69812

Inizan ML, Reduron M, Roche H, Tixier J (1995) Technologie de la pierre taillée. Préhistoire de la pierre taillée (Vol. Tome 4). CREP

Kot M, Tyszkiewicz J, Gryczewska N (2024) Can we read stones? Quantifying the information loss in flintknapping. J Archaeol Sci 161:105905. https://doi.org/10.1016/j.jas.2023.105905

Kot M, Tyszkiewicz J, Leloch M, Gryczewska N, Miller S (2025) Reliability and validity in determining the relative chronology between neighbouring scars on flint artefacts. J Archaeol Sci 175:106156. https://doi.org/10.1016/j.jas.2025.106156

Lin SC, Douglass MJ, Holdaway SJ, Floyd B (2010) The application of 3D laser scanning technology to the assessment of ordinal and mechanical cortex quantification in lithic analysis. J Archaeol Sci 37(4):694–702

Lin SC, Clarkson C, Julianto IMA, Ferdianto A, Jatmiko, Sutikna T (2024) A new method for quantifying flake scar organisation on cores using orientation statistics. J Archaeol Sci 167:105998. https://doi.org/10.1016/j.jas.2024.105998

Linsel F, Bullenkamp JP, Mara H (2024), January 10 Linking Scars: Topology-based Scar Detection and Graph Modeling of Paleolithic Artifacts in 3D. CAA2023 Conference Proceedings. Computer Applications in Archaeology Conference. https://doi.org/10.5281/zenodo.10477448

Linsel F, Bullenkamp JP, Mara H (2025), March 5 From scar to scar: Reconstructing operational sequences of lithic artifacts using scar-ridge-pattern-based graph models. CAA2024 Conference Proceedings. Computer applications in archaeology conference. https://doi.org/10.5281/zenodo.14973253

Liu C (2024) Variation Matters: Expanding the Scope of Experimental Archaeology. Adv Archaeol Pract 12(4):375–389. https://doi.org/10.1017/aap.2024.30

Liu C, Stout D (2023) Inferring cultural reproduction from lithic data: a critical review. Evol Anthropol 32(2):83–99. https://doi.org/10.1002/evan.21964

Lombao D, Falcucci A, Moos E, Peresani M (2023) Unravelling technological behaviors through core reduction intensity. The case of the early Protoaurignacian assemblage from Fumane Cave. J Archaeol Sci 160:105889

Miao Z, Zhang Y, Fabian Z, Hernandez Celis A, Beery S, Li C, Liu Z, Gupta A, Nasir M, Li W, Holmberg J, Palmer M, Gaynor K, Arbelaez P, Wang P, Dodhia R, Ferres JL (2025) New frontiers in artificial intelligence for biodiversity research and conservation with multimodal language models. Methods Ecol Evol. https://doi.org/10.1111/2041-210X.70120

Pargeter J, Brooks A, Douze K, Eren M, Groucutt HS, McNeil J, Mackay A, Ranhorn K, Scerri E, Shaw M, Tryon C, Will M, Leplongeon A (2023) Replicability in Lithic Analysis. Am Antiq 88(2):163–186. https://doi.org/10.1017/aaq.2023.4

Pastoors A, Tafelmaier Y, Weniger G-C (2015) Quantification of late Pleistocene core configurations: Application of the Working Stage Analysis as estimation method for technological behavioural efficiency. Quartär Internationales Jahrbuch Zur Erforschung Des Eiszeitalters Und Der Steinzeit 62:63–84. https://doi.org/10.7485/QU62_3

Porqueddu M-E, Sciuto C, Lamesa A (2023) Reconsidering the Chaîne Opératoire: At the crossroad between people and materials. Open Archaeol. https://doi.org/10.1515/opar-2022-0296

Richardson E, Grosman L, Smilansky U, Werman M (2013) Extracting scar and ridge features from 3D-scanned lithic artifacts. In: Earl G, Sly T, Chrysanthi A, Murrieta-Flores P, Papadopoulos C, Romanowska I, Wheatley D (eds) *Extracting scar and ridge features from 3D-scanned lithic artifacts*. Amsterdam University Press, pp 83–92. https://doi.org/10.1017/9789048519590.010

Richter J (1997) Sesselfelsgrotte III: Der g-schichten-komplex der sesselfelsgrotte zum verständnis des micoquien. ranz Steiner

Sellet F (1993) Chaine Operatoire; The Concept and Its Applications. Lithic Technol 18(1–2):106–112. https://doi.org/10.1080/01977261.1993.11720900

Shipton C, Foulds F, Rawlinson A, Leroyer M, Ashton N, White M (2026) Turning-the-edge, Tranchet, and social signalling at Boxgrove. Camb Archaeol J 36(1):58–75. https://doi.org/10.1017/S0959774325100152

Shott MJ (2007) The Role of Reduction Analysis in Lithic Studies. Lithic Technol 32(1):131–141. https://doi.org/10.1080/01977261.2007.11721048

Soressi M, Geneste J-M (2011) The History and Efficacy of the Chaîne Opératoire Approach to Lithic Analysis: Studying Techniques to Reveal Past Societies in an Evolutionary Perspective. PaleoAnthropology 2011:334–350. https://paleoanthropology.org/ojs/index.php/paleo/article/view/643

Tafelmaier Y, Bataille G, Schmid V, Taller A, Will M (2022) Methods for the Analysis of Stone Artefacts: An Overview. Springer Fachmedien. https://doi.org/10.1007/978-3-658-39091-4

Timbrell L, Scott C, Habte B, Tefera Y, Monod H, Qazzih M, Marais B, Black W, Maroma C, Ndiema E, Henderson S, Elmes K, Plomp K, Grove M (2022) Testing inter-observer error under a collaborative research framework for studying lithic shape variability.

Archaeol Anthropol Sci 14(10):209. https://doi.org/10.1007/s12520-022-01676-2
Tixier J (1963) Typologie de l'Epipaléolithique du Maghreb, vol vol 2. Mémoires du Centre de Recherches anthropologiques et préhistoriques et ethnographiques
Ushey K, Wickham H (2024) Renv: Project environments. https://rstudio.github.io/renv/
Vaesen K, Hussain ST (2026) The perks and perils of replicability. Am Antiq. https://doi.org/10.1017/aaq.2025.10146
Valletta F, Smilansky U, Goring-Morris AN, Grosman L (2020) On measuring the mean edge angle of lithic tools based on 3-D models–a case study from the southern Levantine Epipalaeolithic. Archaeol Anthropol Sci 12(2):49