Archeologia e Calcolatori
31.1, 2020, 55-76
doi 10.19282/ac.31.1.2020.03

A NEW COMPUTATIONAL METHOD TO QUANTIFY MORPHOLOGICAL STANDARDIZATION AND VARIATION WITHIN CERAMIC ASSEMBLAGES

1. Introduction

The appearance of pottery in standard forms and sizes is one strong indicator of mass production and specialization in pre-modern contexts (Rice 1991; Arnold 2000). Quantification of variation within ceramic datasets therefore provides a tool for evaluating scales, organizations, and technological practices behind production and for gauging coordination and complexity of economic systems. Manufacturing of Mediterranean ceramic wares varied widely, from small households and workshops to regionally coordinating centers firing many thousands of consistent pieces (Peacock 1982, 75-128).

Complex and large closed shapes like transport amphoras demanded technical expertise and care but also – in the Roman world – massive outputs and consistency across producers. Archaeological inquiry into standardization of these production systems has often prioritized fabric composition or decoration over form (Kotsonas 2014). Where shape has played a role, it is largely through single-variable measurements like maximum diameter, ideally expressed statistically using the coefficient of variation (CV) (Roux 2003, 772; Orton, Hughes 2013, 147-148). Yet even this approach is uncommon for the Mediterranean, perhaps because high levels of standardization are assumed under such complex economic conditions as those that characterized the Roman Empire. Formal analysis therefore presents opportunities for more systematic evaluation but also clear challenges related to the selection of effective measures (Kvamme *et al.* 1996; Zapassky *et al.* 2006). Thorough characterization of formal features can allow us to quantify finely graduated variations in shape, providing a more analytically rigorous and flexible metric for evaluating the varied attributes linked to scales and systems of production.

This paper presents a new method and computational pipeline to quantify morphological similarities and differences among individual ceramics modeled as point clouds. This data is more available than ever as photogrammetric and other 3D methods are becoming common practice for archaeological documentation, representation, and analysis (e.g., Modrzewska *et al.* 1993, 2010; Bursich, Pace 2017; O'Higgins *et al.* 2019). Projects like ArchAIDE (https://archaide-desktop.inera.it/) have recently invested in ever-expanding knowledge bases of pottery forms, decoration styles, and stamps, integrating these into user-friendly resources like the Archaeology Data Service's "Roman Amphorae: a digital resource" (University of Southampton 2014).

Our method enables comprehensive 3D characterization of geometries down to the pixel level. While many studies have explored, created, or improved methods for generating point clouds, and several have focused on the relative merits of different techniques (Katz, Friess 2014; Evin *et al.* 2016), far less discussion has centered on the archaeological questions, innovative analyses, and new insights they can generate. Various research groups have advocated for total-object comparison (Friess *et al.* 2014; Bevan *et al.* 2014; Hassett, Lewis-Bale 2016), and Spelitz *et al.* (2019) have shown the utility of point clouds for assessing individual ceramic vessels in particular. Yet comparing comprehensively and broadly within and across pottery assemblages demands a quantitative, generalizable, and scalable method. Approaching ceramic shapes as distinct graduated sections ("segmentwise") allows 3D analysis to account for processes of wheel-made ceramics and the varying care afforded to different functional parts of vessels; it also allows the incorporation of partial ceramics, extending the approach to the typically fragmentary material record.

Prior studies have also argued for a computation-heavy approach, demonstrating the analytical rigor achieved by assessing the entire point cloud rather than through more limited landmark-based approaches that focus on specific pre-determined points (Birch, Martinón-Torres 2019). Approaches such as geometric morphometrics can be limited by their requirement for *a priori* knowledge when, in practice, we may not know precisely where to find the most meaningful differences or similarities beforehand. Analyzing full point clouds allows objects to help steer analysis, a possible functional benefit also given interest in automating aspects of typological and stylistic study (Karasik, Smilansky 2011; Bursich, Pace 2017). High-density point clouds increase data-gathering and computation times, but can easily outweigh the costs, as Hassett and Lewis-Bale (2016, 198) note: «While the distribution of points across the surface is not uniform, the redundancy in the sheer number of data points available to be compared in a high-density point cloud addresses concerns that, in order to appropriately compare surfaces, an even sampling strategy must be applied (as, for instance, in semi-landmark approaches)».

Our pipeline was designed for efficiency and scalability. The Python codebase leverages state-of-the-art machine learning algorithms and high-speed data structures for streamlined analysis of large datasets rather than relying on capabilities included in commonly used point cloud visualization software. The modular architecture allows for easy inclusion of open-source algorithms that are constantly improving for speed and performance, with the result that the pipeline can be improved without waiting for new versions of particular software platforms. We have automated away the tedious aspects of processing, relying on human involvement only in situations requiring archaeological judgment. For example, our pipeline auto-detects and auto-denoises using Otsu's algorithm with a single click. Most importantly, the method

relies solely on open-source libraries, ensuring free use for researchers. Free, flexible, powerful, and clean, Python is the programming language of choice across sciences, engineering, and other fields, replacing costly commercial products with expensive add-ons. More technical aspects of the pipeline have been abstracted away to leave a streamlined user-friendly interface. Links to example iPython notebook scripts are included in Section 6.

For analytical demonstration here, we draw on two case studies of Mediterranean amphoras from late antique shipwrecks that provide closed assemblages of jars. Produced in bulk for the packaging of processed agricultural goods, these ceramic forms offer a helpful test case since they aimed at broadly consistent shapes and sizes (Peacock, Williams 1986; Bevan 2014). The methodology, though, should prove useful for other pre-modern wheel-made vessels – and likely other ceramics and even other artifacts – of different forms and contexts.

2. Pipeline

The complex shapes, functions, and corresponding techniques of manufacture associated with different pottery forms demand an approach that is systematic yet flexible. Dividing point clouds into horizontal segments provides an opportunity to characterize wheel-made forms in a way that meaningfully addresses their production processes, and here reflects their assembly from

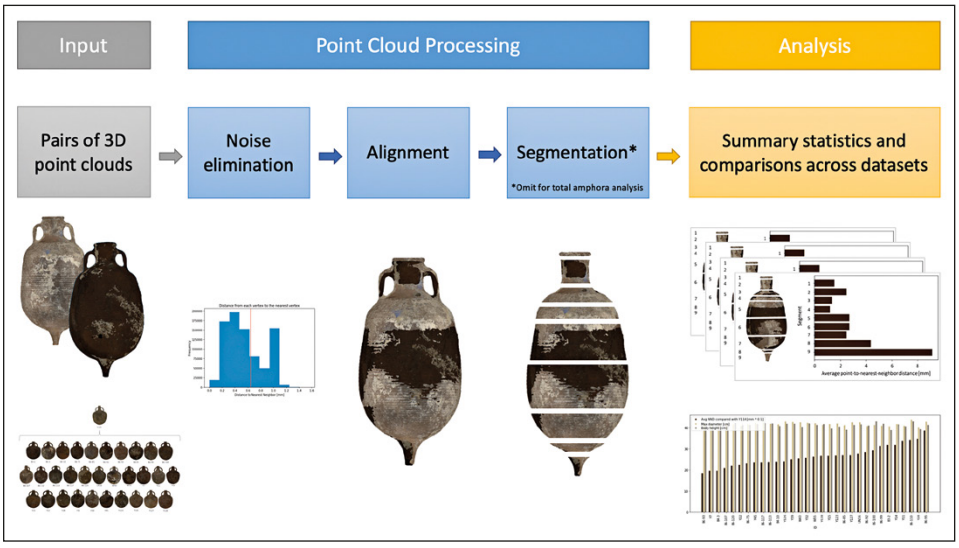

Fig. 1 – The comparative morphological analysis method and pipeline.

separately made parts (e.g., Demesticha 1998). Segmentation allows inclusion of the entire morphology but also emphasis on certain features potentially central to function or other practical concern of the potter or user. In the case of the amphoras here, the various body segments represent the crucial components if volume control was a paramount consideration; some combination of toe, handles, and rim, on the other hand, might be of great importance if stacking for storage or transport, or easy consumer recognition of the origin of the jar's contents was a fundamental concern behind shape (Hein *et al.* 2008; Knapp, Demesticha 2017, 37-41; Lawall 2018). This section explores segments of interest for such bulk transport jars, but the pipeline was designed to meet the varied requirements of different shapes and functions (Fig. 1). In this instance, we rely on point clouds generated by structured light scanning technology using a commercial Artec Eva system, but high-resolution data from other sources could likewise form a reliable basis for analysis.

2.1 *Segmentwise comparison, n = 2*

In this section, segmentwise comparison is described and demonstrated using the point cloud representations of two late Roman amphoras of the Agora M273 type (or Late Roman 8, according to Pieri's typology: Pieri 2005, 132-137) raised from a small and scattered shipwreck likely dating to the 4th century AD. Investigated by the Institute of Nautical Archaeology in 1980, the assemblage lay off the shore near Gümüşlük, the site of ancient Myndos, at the western tip of the Bodrum peninsula in southwest Turkey (Rosloff 1981, 281). This type represents one of a family of forms exhibiting a tall ridged body with low center of gravity, a cylindrical neck with a simple rounded rim and a tapered peg toe. The form evolved over the course of several centuries, from at least the 3rd into the 6th century, at a variety of production centers concentrated in the east Aegean region and along the western shore of Asia Minor (Opaiţ 2014, 443-444). The formal similarities of the two jars selected from the Gümüşlük wreck – 80E-2 and 4634 – are mirrored in their identical fabrics as well as shared inscriptions seemingly indicating for each an intended body volume of 41 units (possibly *xestai*: Pieri 2005, 78-79). The pair therefore provided an ideal case study for developing the proof of concept, allowing investigation of close standardization while also reducing unnecessary variables that could interfere with early pipeline development.

2.1.1 Noise elimination

It is necessary first to remove stray particles and other points corresponding to noise from the 3D recording process that could confound the point cloud analysis. Machine learning is leveraged to detect and remove points that are too far from their neighbors. The basic principle is that a point located far from its nearest points is probably floating in space and therefore not part

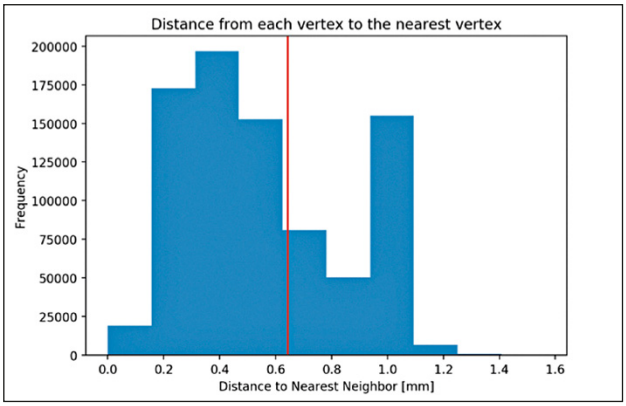

Fig. 2 – Otsu thresholding for outlier detection to eliminate noise. Example: Amphora 4634.

of the surface. Collecting all nearest neighbor distances and plotting them as a histogram reveals two main modes or groupings: those points with close neighbors (Fig. 2, left of the red line) and those far from any neighbors (Fig. 2, right of the red line). The threshold between these two groups is automatically detected using Otsu's method, a well-established machine learning algorithm commonly employed in computer vision and image processing to transform from grayscale to binary images in black and white (Otsu 1979). All points with a nearest neighbor distance greater than the Otsu threshold (Fig. 2, red line) are then removed, leaving behind the clean point cloud surface of the vessel for analysis.

Nearest neighbor distance analysis offers a useful way to clean data captured in the diverse and sometimes challenging conditions of field- and museum-based work, but it also has the potential to introduce a bottleneck in computational efficiency. The high resolution of the structured light scans here is generally beneficial since it means that each amphora point cloud contains hundreds of thousands of points. But for every point, a distance calculation would theoretically be performed with every other point in the cloud. Going beyond typical approaches that rely on distance calculations within existing software packages (e.g., Hassett, Lewis-Bale 2016, who use CloudCompare), our pipeline optimizes query speed and ameliorates runtime by partitioning space and organizing points using the KD-tree data structure as implemented by the scikit-learn machine learning library (Pedregosa *et al.* 2011).

2.1.2 Point cloud alignment and segmenting

Once outliers have been removed, the clean amphora point clouds are then rendered in CloudCompare, an open-source point cloud processing

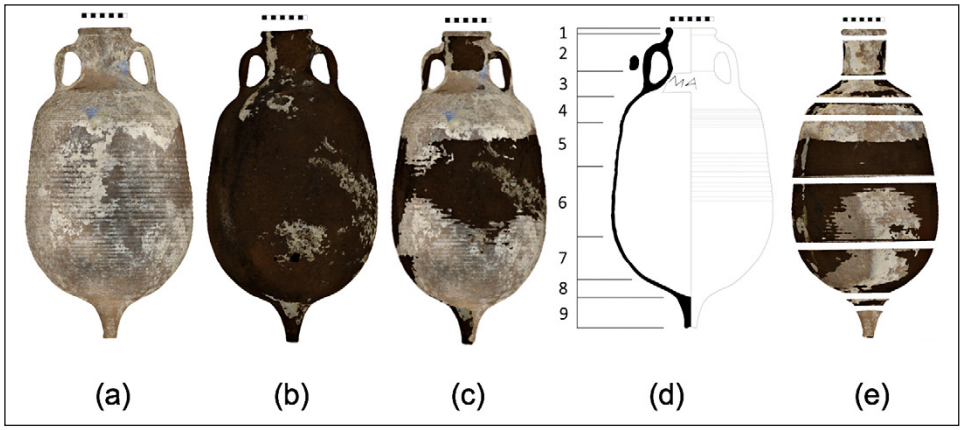

Fig. 3 – The point cloud representations of amphoras a) 80E-2 and b) 4634 from the Gümüşlük site. c) The overlapping point cloud representations of amphoras 80E-2 and 4634, aligned and finely registered. d) Designation for the Gümüşlük amphoras into the following segments: rim (1), neck (2), upper shoulder (3) shoulder (4), upper body (5), middle body (6), lower body (7), bottom (8), and toe (9). Drawing by Bilge Güneşdoğdu. e) The aligned Gümüşlük amphoras segmented into key graduations corresponding to the rim, neck, shoulder, upper body, middle body, lower body, bottom, and toe.

platform (Fig. 3a and 3b) (CloudCompare, version 2.8.1 Gnu Public License software: http://www.cloudcompare.org/). In CloudCompare, the user designates one point cloud as the reference (fixed) object and manually aligns the other (movable) point cloud to maximize overlap. Locations on the movable point cloud are first selected and matched to corresponding points on the fixed point cloud. This step prepares the point clouds for application of the Iterative Closest Point (ICP) algorithm, which finely registers the moving point cloud to the reference point cloud by minimizing the difference between them (Fig. 3c) (Chen, Medioni 1992). Taking advantage of CloudCompare's graphical user interface (GUI), this step also forces the user to analyze visually the point cloud, to conceptualize the expected output, and to observe features – major imperfections, missing pieces, etc. – that could potentially confound analysis. The effect of fine alignment following careful manual alignment is not visually appreciable, but it is consequential to the analysis.

The handles are then removed in this instance. Handles hinder close alignment due to the high variability of their hand-formed shapes and positioning. All isolated handle segments, however, have been stored and future extensions to the pipeline are in development to address their unique attributes, allowing these to play a role in analytical shape comparison where handles are of interest.

Slicing planes then separate the registered point clouds into key horizontal graduations, resulting in distinct segments such as neck, shoulder, lower body, toe, etc. The number and designation of segments (Fig. 3d) is at the user's discretion, based on the questions being asked and the vessel form. Since here we aim to interrogate shape as it relates to the production of containers with not only visually recognizable shapes but also consistent volumes, we have sliced the body at several points where the complex shape changes, and distinguished also the common typological markers of rim, neck, base and toe. CloudCompare's GUI provides support for slicing.

The segments must then be realigned manually, and subsequently registered using ICP. This step promotes segment-specific analogy to prevent, for example, a distortion near the rims from causing the overall alignment to offset the toes, even if the toes are identical in shape. We have found it most practical and effective first to segment out a large chunk containing an area of interest, and then align, trim edges that could confound analysis, and iterate. Maximizing alignment without losing excessive amounts of material depends on user judgment, but the learning curve is not steep. Expertise can be developed after working with just a few models. The resulting point cloud segments (Fig. 3e) are then exported, stored in the Wavefront (OBJ) geometry definition file format, and passed to a script that executes cloud-to-cloud distance analysis on each pair of segments.

2.1.3 Cloud-to-cloud distance analysis

For every segment, the analysis script computes and stores the mean distance from each point in one object to the nearest neighbor in the reference object. With regard to efficiency, the challenges and solutions encountered at this stage are analogous to those involved in outlier detection. The nearest neighbor distance measurements for each segment pair are then pooled and two metrics are computed: the mean nearest neighbor distance from one object to the other summarizes the overlap between the two segments, and the standard deviation quantifies the spread of the distribution across the segment's surface. These metrics can be plotted for each segment to create

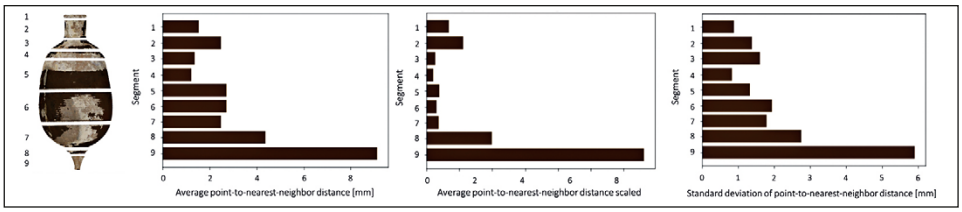

Fig. 4 – Absolute average point-to-nearest-neighbor distance, average point-to-nearest-neighbor distance scaled by the segment's maximum diameter, and standard deviation of point-to-nearest-neighbor distances between amphoras 80E-2 and 4634 for all segments.

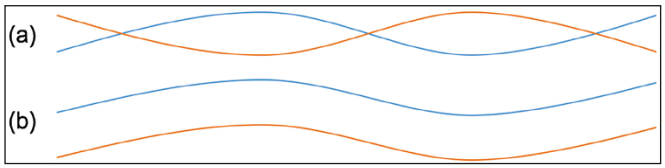

Fig. 5 – Theoretical cross-sectional view of two amphora point surfaces in a segment with a) a large spread in distances between the two surfaces and b) a small spread in distances between the two surfaces.

informative summaries for shape differences between the two objects across the graduated segments (Fig. 4, left plot).

An analysis might proceed down the graduations of Fig. 7, which shows the absolute average point-to-nearest-neighbor distance between amphoras 80E-2 and 4634 for all segments. At the rims (Segment 1) of 80E-2 and 4634, there are about 2 mm of average distance between the two surfaces. The necks (Segment 2) are less uniform. The shoulder segments (Segments 3 and 4) seem to align well. The bodies (Segments 5, 6, and 7) show somewhat more distance from each other. The attachment area at the bottom between the body and toe (Segment 8) aligns poorly, as does the hand-shaped toe (Segment 9). Scaling the average distance between the two point clouds of each segment by the maximum diameter of that segment (Fig. 4, middle plot) offers a better representation of their correspondence given the increasing margin of error associated with the production of larger-diameter parts of these vessels. The resulting histogram shows considerably less variation in the segments comprising and close to the body (Segments 3-7) in comparison with components that were sculpted by hand or adjacent to those components: i.e., Segments 1 (the rim), 2 (where the handles attach, with the attachment points taking much surface area), 8 (where the toe anchors to the amphora body), and 9 (the toe). We should note here that 80E-2 is also missing a chip near the bottom of its toe, accounting for some of the difference measured at Segment 9.

Plotting the standard deviation for point-to-nearest-neighbor distance across all segments reveals patterns in the spread of distances across the amphoras (Fig. 4, right plot). Throughout the lower body segments (excluding the toe), the tight mean distances and relatively larger standard deviations indicate that in these segments, some distances between the amphoras are very small and some distances are larger. This suggests that a randomly selected cross-sectional view of the registered point clouds in these locations might look something like Fig. 5a. On the other hand, a randomly selected cross-sectional view of a segment with a small spread in distances between the two surfaces but the same average distance as in Fig. 5a might look more like Fig. 5b.

2.2 *Segmentwise comparison, n > 2*

This proof of concept sets out a framework for comparing amphoras one-to-one, but translating this method to larger assemblages requires accommodating datasets of size $n > 2$. In the course of point cloud alignment, both manual and computational, one object is naturally designated as the reference object. Building on this idea, the segmentwise approach allows a reference amphora to be compared to any given number of other amphoras.

To explore this analytical framework, another assemblage of amphoras from the same region was selected as an appropriate and larger case study: a set of Aegean Late Roman 2C (Pieri 2005, 89) jars excavated by the Institute of Nautical Archaeology from an important early 7th-century AD shipwreck at Yassıada, again off the tip of the Bodrum peninsula in southwest Turkey (Fig. 6). Likely en route from the east Aegean to the northeast corner of the Mediterranean when it sank in the late 620s, this wreck offers some of our most extensive evidence for a large collection of objects in circulation together at the end of antiquity. The assemblage includes some 900 cargo amphoras (Bass, van Doorninck 1982; van Doorninck 2015), the majority of which belong to this late LR2C type. With its spherical body, cylindrical neck, simple out-turned rim and no toe, this form represents a late derivative of one of the most common and long-lived types of the period, Late Roman 2 (LR2), manufactured throughout the Aegean region since at least the 4th century AD and distributed widely across the Mediterranean, Black Sea, and far beyond (Steckner 1989; Karagiorgou 2001). Thanks to comprehensive excavation in the early 1960s as well as ongoing study of its ceramic assemblage in recent decades, several strongly standardized groups have been identified within the shipwreck assemblage's LR2C jars, of which several hundred survive wholly or mostly intact (van Alfen 2015). One particularly well-defined group that has been the focus of recent work was selected ("spiral-combed" type), providing 32 examples for this study. The form is in preparation for publication, when all data will be made freely available, including not only detailed measurements but the high-resolution point clouds used here.

Designation of the reference object in a dataset of $n > 2$ may be subjective or objective. Different approaches might variously assign the role to the most complete amphora, or the one considered most "average" by some predefined aesthetic or quantitative measure. For this study, amphora Y114 (Fig. 6a) was identified as "average" within this dataset based on a combination of maximum diameter (Fig. 6b) and body height (Fig. 6c). Maximum diameter and body height are demarcated in 6d. Amphoras 86-93, 86-99, W43, and Y128 were then identified as particularly close in their linear dimensions to Y114 (Fig. 6a). Key segments were defined (Fig. 6e), and the process described in Section 2a was executed for each pair as indicated in Fig. 6d. Links to the

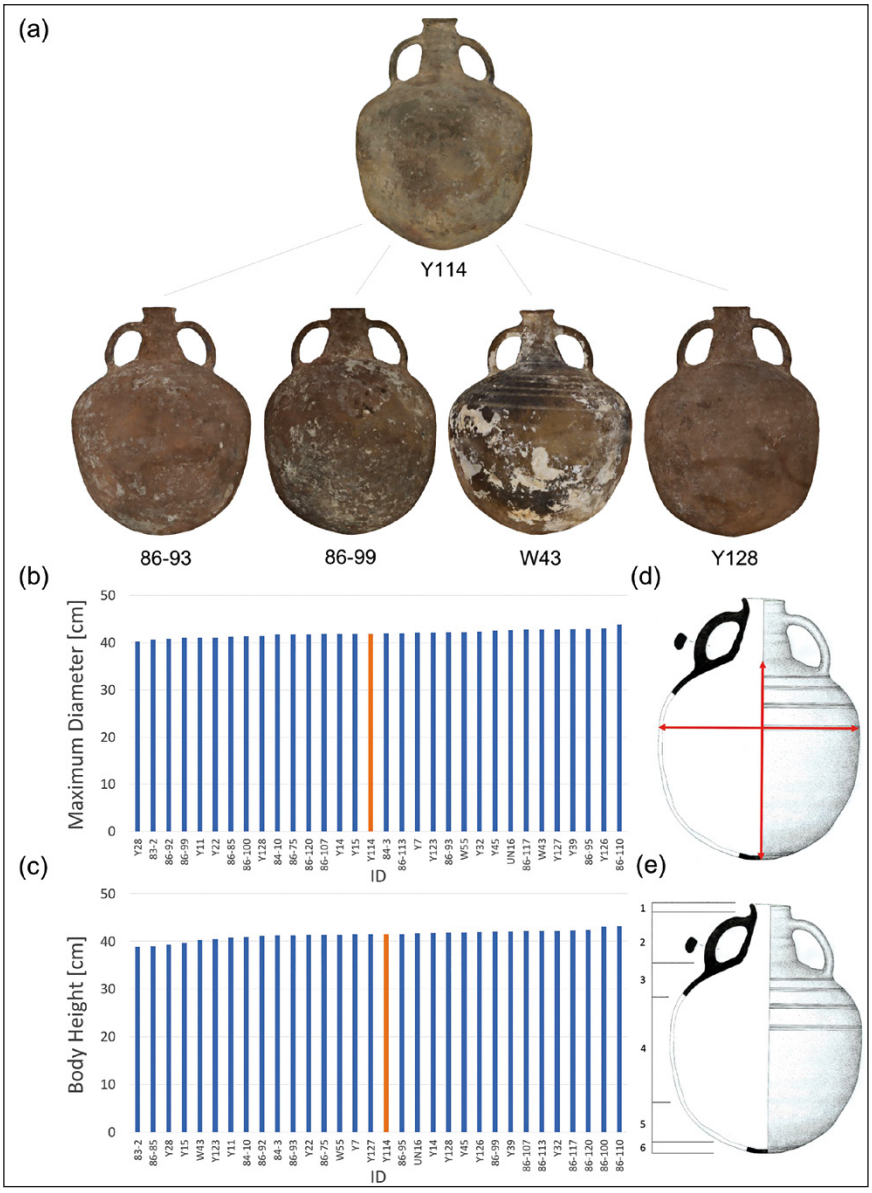

Fig. 6 – a) Point cloud representation of Y114, which was deemed the most average amphora in the "spiral-combed" group from the Yassıada assemblage. Members of the sample set for segmentwise comparison, n > 2. 86-93, 86-99, W43, and Y128, are particularly close in dimensions to Y114. b) Maximum diameters within the dataset. c) Body heights within the dataset. d) Demarcations for maximum diameter and body height for the jars. Amphora drawing by Seçil Kayacık. (e) Segmenting an amphora at the rim (1), neck (2), shoulder (3), main body (4), lower body (5), and base (6).

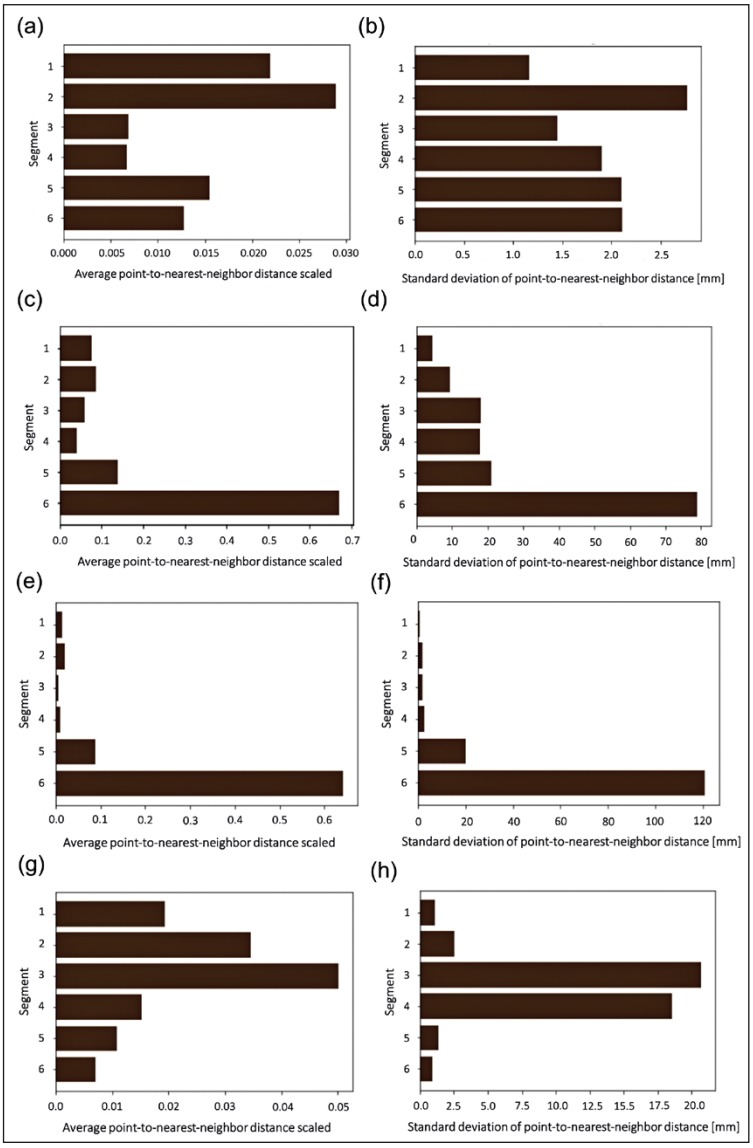

Fig. 7 – a-b) Summary statistics for the segmentwise comparison of 86-93 to Y114: a) average point-to-nearest-neighbor distance scaled, and b) standard deviation of point-to-nearest-neighbor-distance. c-d) Summary statistics for the segmentwise comparison of 86-99 to Y114: c) average point-to-nearest-neighbor distance scaled, and d) standard deviation of point-to-nearest-neighbor-distance. e-f) Summary statistics for the segmentwise comparison of Y114 to W43: e) average point-to-nearest-neighbor distance scaled, and f) standard deviation of point-to-nearest-neighbor-distance. g-h) Summary statistics for the segmentwise comparison of Y114 to Y128: g) average point-to-nearest-neighbor distance scaled, and h) standard deviation of point-to-nearest-neighbor-distance.

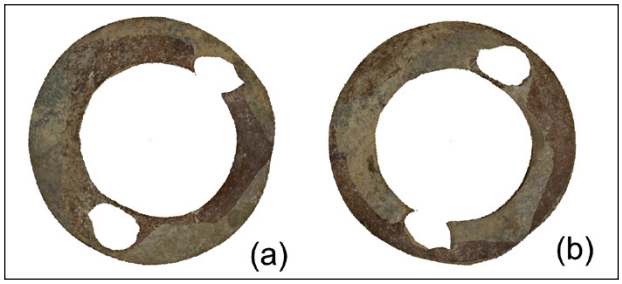

Fig. 8 – a) Top and b) bottom views of the aligned point clouds for the shoulders of amphoras Y128 and Y114 reveal regions where large distances could have accumulated, resulting in the marked spike around Segment 3 on the plots shown in Fig. 7g, h.

iPython notebooks are included in Section 6. In what follows, we summarize the results for the segmentwise comparison of each pair in this dataset.

The plot for average point-to-nearest-neighbor distance scaled for the segmentwise comparison of 86-93 to Y114 (Fig. 7a, b) echoes the distribution for the Gümüşlük amphoras: larger distances in the neck and smaller distances in the body region. A similar rationalization may apply to these Yassıada jars: the rim (Segment 1) is somewhat hand-sculpted but may have been more easily measured and controlled than the neck (Segment 2). The two upper body segments (3 and 4) are well-controlled, and the lowest body segment (5) may have experienced some deformation from the weight during drying. The most notable difference between this pair of jars and those from Gümüşlük is in the last (lowest) segment of each distribution, where the Yassıada amphoras have flatter bases which align more tightly than the Gümüşlük toes. As with the Gümüşlük amphoras, the standard deviation of distances is greatest at the neck (Segment 2), and then drops at the top of the body (Segment 3) before increasing toward the base.

The segmentwise comparison of 86-99 to Y114 (Fig. 7c, d) yields a similar distribution to those from Gümüşlük and 86-93 above. The distances overall, however, are much greater, especially at the bases. This comparison of 86-99 to Y114 illustrates that bases can have varying degrees of concavity. The distributions of summary statistics for the segmentwise comparison of W43 to Y114 (Fig. 7e, f) are congruent with those for the comparisons of both 86-93 and 86-99 to Y114, with the exception of the two lowest segments, especially the base. The markedly different base shapes of W43 and Y114 align poorly. The comparison distributions for Y128 to Y114 (Fig. 7g, h) would again be similar to those from Gümüşlük and 86-93 to Y114 if not for the small differences in the lower body and toe sections and, most notably, the spike in distance at the shoulder (Segment 3). The point clouds for the

shoulder segment indicate that there are regions where large distances could have accumulated even after tight alignment (Fig. 8).

In analyzing these distributions, it is important to remember that the units given on these plots are in mm. Considering this scale, the results seem to suggest overall that the producers of these jars were able to control and standardize shape to an exacting degree, especially where it mattered most for body volume. This comparison framework allows detailed evaluation of an entire set of amphoras and the ways their individual components deviate morphologically from those of some selected reference.

3. Total amphora comparison, $N \geq 2$

The analysis pipeline from Section 2 also makes it possible to characterize a larger set of jars by morphological deviation from a reference jar. This approach follows in the tradition of using single measurements like height, maximum diameter, or volume, but offers a more holistic and nuanced formal analysis. While proof-of-concept methods papers need not always rely on statistically significant datasets, especially given the limited availability in many archaeological contexts (Evin *et al.* 2016), we were nonetheless able to

Fig. 9 – The sample set of 32 jars selected for evaluation of morphological deviation.

ID	Max. Diam.	Body Height
83-2	40.7	38.9
84-3	42.0	41.3
84-10	41.8	40.9
86-75	41.8	41.4
86-85	41.3	39.0
86-92	40.9	41.2
86-93	42.3	41.3
86-95	43.0	41.5
86-99	41.1	42.1
86-100	41.4	43.1
86-107	41.9	42.2

ID	Max. Diam.	Body Height
86-110	43.9	43.2
86-113	42.0	42.2
86-117	42.8	42.3
86-120	41.8	42.4
UN16	42.7	41.7
W43	42.8	40.3
W55	42.3	41.4
Y7	42.2	41.5
Y11	41.1	40.8
Y14	41.9	41.8
Y15	41.9	39.7

ID	Max. Diam.	Body Height
Y22	41.1	41.4
Y28	40.2	39.3
Y32	42.4	42.2
Y39	42.9	42.1
Y45	42.6	41.9
Y114	41.9	41.5
Y123	42.2	40.5
Y126	43.1	42.0
Y127	42.8	41.5
Y128	41.5	41.9

Tab. 1 – Physical measurements in a table format to be passed into the analysis pipeline. The columns correspond to object ID, maximum diameter, and body height.

Morphological Difference from Y114			
	Mean	SD	
NND	0.0485 [cm]	0.0233 [cm]	
Physical Metrics, Absolute			
	Mean	SD	CV
Max. diam.	42.31 [cm]	0.93 [cm]	2.20
Body height	41.53 [cm]	0.95 [cm]	2.29

Tab. 2 – Mean, standard deviation, and coefficient of variation of morphological deviation from the designated jar across the dataset, reported alongside the same summary statistics for physically measured traditional metrics.

identify and record 32 amphoras that were sufficiently intact and typologically uniform to support a larger-scale analysis.

To evaluate the average, spread, and variation of morphological deviation across this dataset, each jar is again compared to a designated reference on the basis of a consistent single overall metric: in this case the average cloud-to-cloud nearest neighbor distance over the entire jar. Continuing with the same ("spiral-combed") group of jars from the Yassıada shipwreck, Y114 was kept as the reference jar to maintain consistency and to allow comparability of our results as analysis scaled up. Additional amphoras were drawn from those morphologically similar – i.e., close in measurable dimensions – to Y114, and also completely intact and accessible for 3D recording (Fig. 9).

The most basic physical measurements (maximum diameter and body height) that had already been manually acquired for each of these amphoras were encoded in a comma-separated value (CSV) file (Tab. 1). Based on these measurements, certain possible "outliers" were identified prior to analysis, which could then be compared with the results from the overall morphological views of these jars. For each pair (reference jar and comparison jar), a Python

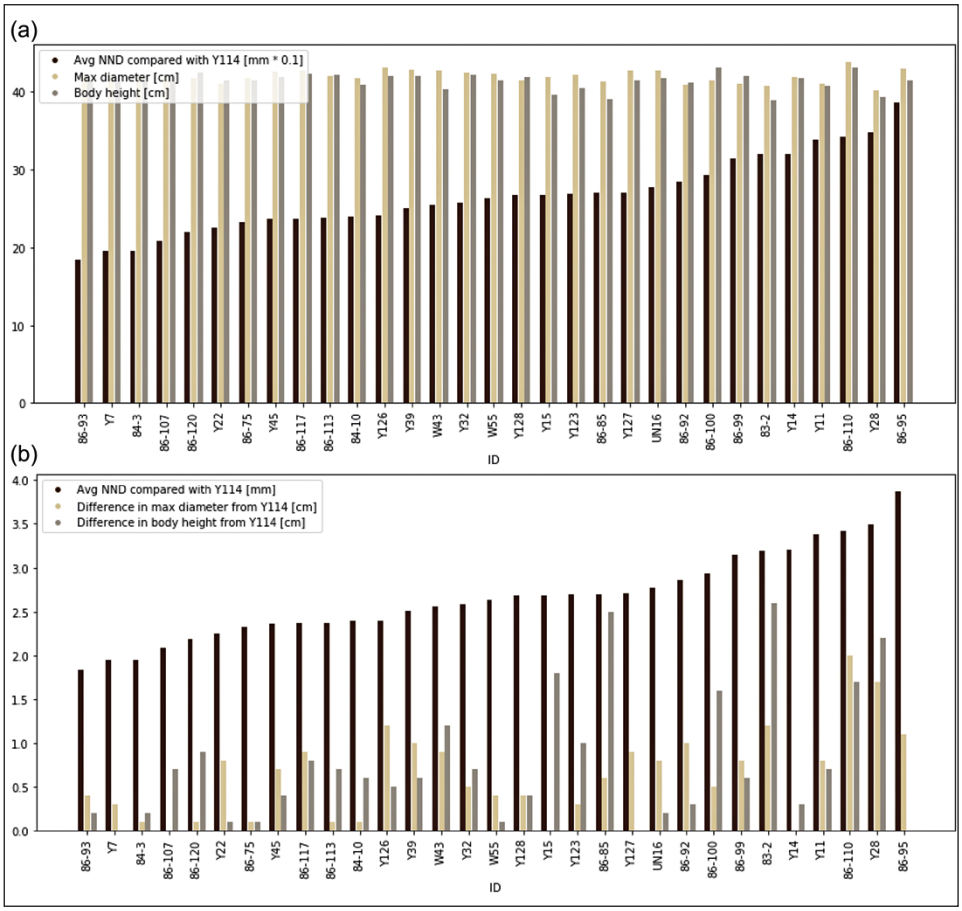

Fig. 10 – a) Morphological difference (measured as average nearest neighbor distance) from Y114 and absolute physical measurements. Note that these average nearest neighbor differences are shown with an alternate vertical axis scale (tenths of mm) to allow easier comparison with the dimensional differences. b) Morphological difference (average nearest neighbor distance) from Y114 compared with differences in physical measurements. Note that the y-axis has been scaled to different units (mm vs. cm) to allow easier comparison.

script executed the process described in Section 2a using the entire amphora point cloud without segmentation. Links to the iPython notebooks for this analysis are included in Section 6.

Overall trends in shape differences and physical measurements were illuminated by several analytics added to the pipeline. The CSV file containing physical measurements was then passed into the analysis pipeline to generate one plot comparing morphological differences with the reference amphora and

absolute physical measurements for every member of the dataset (Fig. 10a), and another plot comparing morphological and physical measurement differences with the reference amphora for every member of the dataset (Fig. 10b).

The greatest shape differences correspond to those amphoras with larger differences in maximum diameter and/or height, including 86-95 and Y28. A significant difference in shape is also evident for amphora 86-110, which initial observations had marked as somewhat atypical in linear dimensions. Smaller shape differences appear for amphoras 86-93, Y7, and 84-3, which all have quite similar dimensions to Y114. One might expect that overall morphological deviation would correlate strongly to deviation in physical dimensions, but Fig. 10b illustrates that the two do not run wholly parallel in all instances. Assuming maximum diameter and height were key parameters guiding the manufacture of amphoras, potters seem to have found ways to control capacities with sufficient consistency regardless of whether those jars ended up exhibiting slightly higher or lower morphological deviations.

The final step of the pipeline calculates the mean and standard deviation of these morphological differences from the designated jar across the dataset, reporting these values alongside the same summary statistics for traditional physical measurements and their differences from the analogous metrics of the designated comparison jar (Tab. 2). These statistics reveal that differences in linear dimensions – and by extension volumes –mong jars in this overall group were controlled within an impressively narrow range, with CV values under 3 indicating minimal variation in comparison with traditional ceramic workshops and even those operating on quite large scales (Roux 2003, 776-780).

4. Pipeline evaluation

Previous sections have touched upon efficiency, and at present the pipeline is well-optimized. As indicated in Section 2a, certain aspects of object alignment and segmenting in CloudCompare must be carried out manually; performing these operations automatically without user supervision could lead to unexpected behaviors with origins that would be difficult to track. The alignment and segmentation of each object pair varies according to the user experience: about 30 minutes each for the first few pairs, and then 10-15 minutes per jar thereafter. Carrying out the required analysis operations upon hundreds of thousands of 3D data points results in total runtimes on the order of minutes on a 2.7-GHz processor (i.e., the built-in processor of a typical Apple MacBook Pro). The user interface generates status messages to indicate whether the analysis is progressing smoothly and thereby mitigate wasted runtime. A strength of the pipeline is that it works directly with the point cloud without having to fit a mesh to the surface, a computationally intensive step which can introduce errors (Memoli, Sapiro 2004).

5. Conclusions

The general methodology and pipeline represent a promising path toward more systematic quantification of formal variation within archaeological datasets. These metrics offer a new opportunity to explore scales and systems of production, and to identify possible patterns and mechanisms of standardization. While the main focus here has been on demonstrating the robustness of the approach, some first conclusions can already be drawn regarding the serial production of at least two groups of late antique Aegean amphoras. The ability to evaluate differences across the entire amphora reflects the most important innovation, allowing us to incorporate finer details of shape than previously possible through single measurements like maximum diameter or body height or landmark-based approaches that have guided most work to date. Certain parts of the jar can be shown to have received greater and lesser attention, enabling archaeologists to track priorities and strategies in production across space and time. In the case studies here, for example, Aegean potters controlled shapes to a remarkable degree in general. The body region that dictated volume seems often to have taken on paramount importance, in the case of the Gümüşlük amphoras showing stronger correspondence than any other part. The analysis reveals evidence for how set capacities were ensured for jars embedded in bulk transport economies, and also how these efforts could be compromised during production, as here through what appears to be deformation in the lower bodies that reduced heights of some Yassıada jars. Basic measurements of height and maximum diameter hint at such controls, but analysis over the entire form help us track how, and evaluate how successfully, these linear dimensions were translated into consistent shapes (e.g., Zapassky *et al.* 2006; Finkelstein *et al.* 2011).

Ongoing work centers on extending the pipeline to a broader range of subjects and analytical capabilities. As discussed throughout Section 2, the current pipeline handles only intact amphoras. Several conceptual questions arise with amphora datasets lacking some or most of their components. Many other Yassıada jars, for example, are limited to the neck, handles, and shoulders, which reflect generally the most diagnostic parts of the amphoras and therefore could still provide a useful analytical sample. Most pertinent is the question of how large the rough overlap between two fragments would have to be for meaningful shape comparison. The analysis also currently excludes handles because variability in their hand-formed shapes and positioning hinder close alignment. Since handles are not part of the volume-related portion of the amphora, though, one might hypothesize that controlling their shape and size was a lower priority. Yet certain of their details may have allowed distinction (during production or transport) among the jars of different workshops or merchants. Should this demand systematic attention, further work could establish a firm basis for comparison since visual inspection reveals greater variation than among most other features. It might be fruitful also to analyze

them separately from the rest of the amphora and think about a framework of "handle microstyle" – in the way Bevan *et al.* (2014) analyzed the "ear microstyle" of Qinshihuang's terracotta warriors – where segmenting handles into the main part and attachment areas offers one possible approach.

Future work aims to explore other closely defined groups within this massive dataset of 900 Yassıada amphoras as well as those from select other assemblages for which such standardization seems less prevalent or precise, a crucial step toward establishing baselines for what constitutes "standardization" and grounding our findings within the broader context of ancient productive systems and the containerization phenomenon (Bevan 2014; Shryock, Smail 2018). We are also developing functionality to evaluate the circularity of the largest cross-sectional slice, that is, to calculate how the amphora's surface at the graduation of greatest diameter deviates from an ideal circle. Varying levels of circularity could provide additional insights into standardization and the practicalities or limits of ancient systems aiming at containers with consistent volumes. But alongside improving our approach to certain Mediterranean closed vessel forms in this way, we aim to employ the methodology on other ceramic and also non-ceramic forms for which such metrics might provide new evidence of ancient productive systems not captured through simple linear dimensions: for example, the production of complex carved architectural elements, figurines, or other sculptural forms (Asgari 1995). Although 3D recording of individual objects is increasingly a concern of many projects, the technologies and infrastructures used to capture this data vary widely, so we are also exploring the parameters for best translating this methodology to other popular 3D data sources like rapid low-cost structure from motion / multi-view stereo (SFM-MVS) and photogrammetry, which have been demonstrated as comparable to structured light scanning in utility for analysis, albeit in a lower resolution (Katz, Friess 2014; Bevan *et al.* 2014). With each improvement and extension to the methodology, the codebase will be updated accordingly and all changes will be backward-compatible.

6. Links to Notebooks

Segmentwise analysis, Gümüşlük amphoras

80E2 and 4634

https://nbviewer.jupyter.org/github/vophamhi/Amphora/blob/master/gumusluk/compare_segmentwise_gumusluk.ipynb

Segmentwise analysis, Yassıada amphoras

Y114 and 86-93

https://nbviewer.jupyter.org/github/vophamhi/Amphora/blob/master/yassiada/Y114_86-93/compare_segmentwise_Y114_86-93.ipynb

Y114 and 86-99

https://nbviewer.jupyter.org/github/vophamhi/Amphora/blob/master/yassiada/Y114_86-99/compare_segmentwise_Y114_86-99.ipynb

Y114 and W43
https://nbviewer.jupyter.org/github/vophamhi/Amphora/blob/master/yassiada/Y114_W43/compare_segmentwise_Y114_W43.ipynb

Y114 and Y128
https://nbviewer.jupyter.org/github/vophamhi/Amphora/blob/master/yassiada/Y114_Y128/compare_segmentwise_Y114_Y128.ipynb

Total amphora comparison, Yassıada amphoras

Full Amphora Comparative Morphological Analysis Part 1: Outlier Removal
https://nbviewer.jupyter.org/github/vophamhi/amphora/blob/master/yassiada/total_amphora/part_1_remove_outliers.ipynb

Full Amphora Comparative Morphological Analysis Part 2: Analysis
https://nbviewer.jupyter.org/github/vophamhi/amphora/blob/master/yassiada/total_amphora/part_2_compare_full.ipynb

Jenny Vo-Phamhi
Stanford Archaeology Center
Stanford University, California
jennyvo@stanford.edu

Justin Leidwanger
Department of Classics
Stanford University, California
jleidwa@stanford.edu

Acknowledgements

We thank the Turkish Ministry of Culture and Tourism for permission to undertake long-term study of this material, and Frederick van Doorninck, Jr. and George Bass for their support of research into the Yassıada shipwreck and survey amphoras. The Bodrum Museum of Underwater Archaeology and the Institute of Nautical Archaeology's Bodrum Research Center provided key support, where particular thanks are owed to Tayfun Selçuk and Hande Savaş, and tao Tuba Ekmekçi Littlefield and Esra Altınanıt Bicer. The Institute of Nautical Archaeology's van Doorninck Endowment helped fund 3D data collection and initial processing, which were undertaken in the field by James Gross, Esther Knegt, and Nicole Gavin. Elizabeth Vo-Phamhi assisted with computational analysis. Special thanks to James Gross for discussion of the jars, their features, and recording processes.

REFERENCES

Arnold D.E. 2000, *Does the standardization of ceramic pastes really mean specialization*? «Journal of Archaeological Method and Theory», 7, 4, 333-375.

Asgari N. 1995, *The Proconnesian production of architectural elements in Late Antiquity, based on evidence from the marble quarries*, in C. Mango, G. Dagron (eds.), *Constantinople and its Hinterland: Papers from the Twenty-Seventh Spring Symposium of Byzantine Studies (Oxford 1993)*, Aldershot, Variorum, 263-288.

Bass G.F., van Doorninck Jr. F.H. (eds.) 1982, *Yassi Ada I: A Seventh-Century Byzantine Shipwreck*, College Station, Texas A&M University Press.

Bevan A. 2014, *Mediterranean containerization*, «Current Anthropology», 55, 4, 387-418.

Bevan A. *et al.* 2014, *Computer vision, archaeological classification and China's terracotta warriors*, «Journal of Archaeological Science», 49, 249-254.

Birch T., Martinón-Torres M. 2019, *Shape as a measure of weapon standardisation: From metric to geometric morphometric analysis of the Iron Age 'Havor' lance from Southern Scandinavia*, «Journal of Archaeological Science», 101, 34-51.

Bursich D., Pace A. 2017, *Ripensando il "metodo Beazley". Ceramica attica e fotomodellazione 3D: il caso del Painter of Syracuse 19861*, «Archeologia e Calcolatori», 28.1, 73-91 (https://doi.org/10.19282/AC.28.1.2017.05).

Carlson D.N., Leidwanger J., Kampbell S.M. (eds.) 2015, *Maritime Studies in the Wake of the Byzantine Shipwreck at Yassıada, Turkey*, College Station, Texas A&M University Press.

Chen Y., Medioni G. 1992, *Object modelling by registration of multiple range images*, «Image and Vision Computing», 10, 3, 145-155.

Demesticha S. 1998, *Experimenting on amphora manufacture*, in E. Karpodini-Dimitriadi (ed.), *Ethnography of European Traditional Cultures: Arts, Crafts, Techniques of Heritage*, Athens, Institute of Cultural Studies of Europe and the Mediterranean, 139-149.

Evin A. *et al.* 2016, *The use of close-range photogrammetry in zooarchaeology: Creating accurate 3D models of wolf crania to study dog domestication*. «Journal of Archaeological Science: Reports», 9, 87-93.

Finkelstein I. *et al.* 2011, *Phoenician 'torpedo' amphoras and Egypt: Standardization of volume based on linear dimensions*, «Ägypten und Levante», 21, 249-259.

Hassett B.R., Lewis-Bale T. 2016, *Comparison of 3D landmark and 3D dense cloud approaches to hominin mandible morphometrics using Structure-from-Motion*, «Archaeometry», 59, 1, 191-203.

Hein A. *et al.* 2008, *Koan amphorae from Halasarna – Investigations in a hellenistic amphora production centre*, «Journal of Archaeological Science», 35, 1049-1061.

Karagiorgou O. 2001, *LR2: A container for the military annona on the Danubian border?*, in S. Kingsley, M. Decker (eds.), *Economy and Exchange in the East Mediterranean during Late Antiquity*, Oxford, Oxbow Books, 129-166.

Karasik A., Smilansky U. 2011, *Computerized morphological classification of ceramics*, «Journal of Archaeological Science», 38, 10, 2644-2657.

Katz D., Friess M. 2014, *Technical Note: 3D from standard digital photography of human crania. A preliminary assessment*, «American Journal of Physical Anthropology», 154, 152-158.

Knapp A.B., Demesticha S. 2017, *Mediterranean Connections: Maritime Transport Containers and Seaborne Trade in the Bronze and Early Iron Ages*, London, Routledge.

Kotsonas A. 2014, *Standardization, variation, and the study of ceramics in the Mediterranean and beyond*, in A. Kotsonas (ed.), *Understanding Standardization and Variation in Mediterranean Ceramics: Mid 2nd to Late 1st Millennium BC*, Leuven, Peeters, 7-23.

Kvamme K.L., Stark M.T., Longacre W.A. 1996, *Alternative procedures for assessing standardization in ceramic assemblages*, «American Antiquity», 61, 1, 116-126.

Lawall M. 2018, *Regional styles of transport amphora production in the Archaic Aegean*, in S. Handberg, A. Gadolou (eds.), *Material Koinai in the Greek Early Iron Age and Archaic Period. Acts of an International Conference at the Danish Institute at Athens*, Monographs of the Danish Institute at Athens 22, Aarhus, Aarhus University Press, 289-311.

Memoli F., Sapiro G. 2004, *Comparing point clouds*, in *Second Eurographics Symposium on Geometry Processing (Nice 2004)*, 32-40 (dl.acm.org/doi/10.1145/1057432.1057436).

Modrzewska I., Oddone M., Pianetti F., Taroni G. 1993, *Anfore spagnole nel Veneto. Progetto di ricerca e primi prodotti*, «Archeologia e Calcolatori», 4, 127-135.

Modrzewska I., Taroni G., Pianetti F. 2010, *Un'anfora frammentaria dalla Laguna di Venezia*, «Archeologia e Calcolatori», 21, 201-210.

O'Higgins P., Fitton L.C., Godinho R.M. 2019, *Geometric morphometrics and finite elements analysis: Assessing the functional implications of differences in craniofacial form in the hominin fossil record*, «Journal of Archaeological Science», 101, 159-168.

Opaiţ A. 2014, *The baggy amphora shape: A new fashion?*, in N. Poulou-Papadimitriou, E. Nodarou, V. Kilikoglou (eds.), *LRCW 4. Late Roman Coarse Wares, Cooking Wares and Amphorae in the Mediterranean. Archaeology and Archaeometry*, BAR International Series 2616, Oxford, Archaeopress, 441-450.

Orton C., Hughes M. 2013, *Pottery in Archaeology*, 2nd Ed., Cambridge, Cambridge University Press.

Otsu N. 1979, *A threshold selection method from gray-level histograms*, «IEEE Transactions on Systems, Man and Cybernetics», 9, 1, 62-66.

Peacock D.P.S. 1982, *Pottery in the Roman World: An Ethnoarchaeological Approach*, London, Longman.

Peacock D.P.S., Williams D.F. 1986, *Amphorae and the Roman Economy: An Introductory Guide*, London, Longman.

Pedregosa F. *et al.* 2011, *Scikit-learn: Machine learning in Python,* «Journal of Machine Learning Research», 12, 2825-2830.

Pieri D. 2005, *Le commerce du vin oriental à l'époque byzantine, Ve-VIIe siècles: le témoignage des amphores en Gaule*, Beirut, Institut Français du Proche-Orient.

Rice P.M. 1991, *Specialization, standardization, and diversity: A retrospective*, in R.L. Bishop, F.W. Lange (eds.), *The Ceramic Legacy of Anna O. Shepard*, Niwot, University Press of Colorado, 257-279.

Rosloff J.P. 1981, *INA's 1980 Turkish underwater survey*, «International Journal of Nautical Archaeology», 10, 4, 277-286.

Roux V. 2003, *Ceramic standardization and intensity of production: Quantifying degrees of specialization*, «American Antiquity», 68, 4, 768-782.

Shryock A., Smail D.L. 2018, *On containers: A forum*, «History and Anthropology», 29, 1, 1-6.

Spelitz S., de Almeida V.M., Lang-Auinger C. 2019, *Automatic geometry, metrology, and visualisation techniques for 3D scanned vessels*, «Digital Applications in Archaeology and Cultural Heritage» (https://doi.org/10.1016/j.daach.2019.e00105).

Steckner C. 1989, *Les amphores LR 1 et LR 2 en relation avec le pressoir du complexe ecclésiastique des Thermes de Samos. Relation entre forme et fonction d'un vase en tant que récipient d'un contenu spécifique. Examen par ordinateur et expérimentation*, in V. Déroche, J.-M. Speiser (eds.), *Recherches sur la céramique byzantine*, «Bulletin de Correspondance Hellénique», Suppl. 18, 57-71.

University of Southampton 2014, *Roman Amphorae: A Digital Resource*, York, Archaeology Data Service (https://doi.org/10.5284/1028192).

van Alfen P.G. 2015, *The restudy of the LR2 amphoras from the seventh-century Yassıada shipwreck: Preliminary evidence for standardization*, in Carlson, Leidwanger, Kampbell 2015, 17-34.

van Doorninck Jr. F.H. 2015, *The seventh-century Byzantine ship at Yassıada and her final voyage: Present thoughts*, in Carlson, Leidwanger, Kampbell 2015, 205-216.

Zapassky E., Finkelstein I., Benenson I. 2006, *Ancient standards of volume: Negevite Iron Age pottery (Israel) as a case study in 3D modeling*, «Journal of Archaeological Science», 33, 1734-1743.

ABSTRACT

Analysis of ceramic standardization and variation provides a powerful tool for evaluating the scale, organization, and technological practices behind pre-modern production and for gauging the coordination and complexity of past economic systems. The selection of formal attributes to allow effective measurement and comparison of complex shapes,

though, presents a crucial challenge to systematic study. Alongside fabric composition and surface treatment, consistent linear dimensions offer helpful metrics for assessing standardized production. More difficult to measure, though, are the many finely graduated variations in shape that can reflect how these processes were implemented and the limits to large-scale serial productions like those of the ancient Mediterranean world. We offer here a new method and computational pipeline, developed using open-source libraries, to quantify morphological similarities and differences among ceramics. Grounded in point cloud comparison, our method enables comprehensive 3D characterization of geometries down to the pixel level and leverages state-of-the-art machine learning algorithms and high-speed data structures for efficiency and scalability across large assemblages. Case studies of transport amphoras from two late antique shipwrecks off the coast of southwest Turkey demonstrate the robustness of the methodology and pipeline. Together, they provide an analytically rigorous and flexible approach to quantifying formal variation within a dataset. The first results suggest strategies for controlling the capacities of these transport jars within late ancient systems of production, but the method should also prove useful in formal analysis of artifacts of other forms and contexts.