



# Shape as a measure of weapon standardisation: From metric to geometric morphometric analysis of the Iron Age ‘Havor’ lance from Southern Scandinavia



Thomas Birch<sup>a,\*</sup>, Marcos Martinón-Torres<sup>b</sup>

<sup>a</sup> Centre for Urban Network Evolutions (UrbNet), Aarhus University, Moesgård Allé 20, 4230, 2nd Floor, DK-8270, Højbjerg, Denmark

<sup>b</sup> Department of Archaeology, University of Cambridge, Downing Street Cambridge, CB2 3ER, United Kingdom

## ARTICLE INFO

**Keywords:**  
Standardisation  
Geometric morphometrics  
Shape analysis  
Variation  
Weapon  
War booty  
Symmetry

## ABSTRACT

Traditionally, standardisation of manufacture has been investigated using metrics (e.g. length and width) and compared in terms of the coefficient of variation (CV). This paper argues that standardisation should not only be investigated via metrics, but also in terms of shape. An Iron Age lance head type ('Havor'), known from three main weapon depositions in Southern Scandinavia, is used as a case study to test the effectiveness of shape analysis against traditional metric analysis for investigating standardisation. Geometric morphometric (GMM) analysis is used to measure the overall shape variation and to test shape difference of the same lance type coming from three different archaeological sites. The results demonstrate GMM to complement the traditional metric approach. Whilst metric measurements offer insights into Havor lance standardisation, the results from multivariate analysis of GMM data provides further explanation about the source of variation in terms of shape, including an assessment of object symmetry. This paper represents the first known methodological application of GMM analysis to iron weapons and demonstrates it to be an effective method for studying product standardisation in terms of shape variation.

## 1. Introduction

The aim of this study is to establish whether shape analysis, through geometric morphometrics, can provide more information about object standardisation than traditional techniques. The paper begins with a short background into the analysis and archaeological study of standardisation, highlighting the need to test and explore shape analysis as a new method of investigation. It then follows to present the results of a study focused on so-called 'Havor' lances – Iron Age weapons from Southern Scandinavia. We present firstly a traditional metric analysis, which is further developed to include product-moment correlations (pair-wise study of variables), in an effort to avoid studying specific features or dimensions in isolation. A GMM analysis is then conducted, with the results subsequently compared to the preceding metric approach using multivariate methods.

### 1.1. Traditional metric analysis: coefficient of variation (CV)

Eerkens and Bettinger define standardisation as a “relative measure of the degree to which artifacts are made to be the same” (2001, p.

493). The study of standardisation in object features such as size, shape or composition can be a useful approach to craft production. For example, high degrees of standardisation may be taken as indicative of a relatively small number of producers, rigid quality control, or high manufacturing skill – among other possibilities. A challenge for standardisation studies is that this is a relative parameter: for instance, the difficulties in establishing thresholds for determining ‘high’ vs ‘low’ standardisation based on the variability of metallurgical slag have been noted (Humphris et al., 2009). In addition, comparing degrees of standardisation can be risky when operating at variable scales or with different technologies, i.e. objects produced by casting in moulds vs those made by sculpting or modelling.

The coefficient of variation is used increasingly as a dimensionless measure of standardisation, as it effectively allows comparisons between datasets with different means, such as the width and length of artefacts (cf. Eerkens, 2000; Eerkens and Bettinger, 2001; Eerkens and Lipo, 2005; Martinón-Torres et al., 2014; Roux, 2003; Stark, 1995, and references therein). The CV is calculated and expressed as a percentage:

\* Corresponding author.

E-mail addresses: [t.birch@cas.au.dk](mailto:t.birch@cas.au.dk) (T. Birch), [m.martinon-torres@arch.cam.ac.uk](mailto:m.martinon-torres@arch.cam.ac.uk) (M. Martinón-Torres).

$$CV = \left( \frac{\text{standard deviation}}{\text{mean}} \right) \times 100$$

Lower CVs correspond with higher standardisation, and vice-versa. When considering manual production systems where moulds or measuring aids are not used, limitations in human perception provide the lower threshold of variability that can be expected, which in turn denotes the highest degree of standardisation. Eerkens (2000; see also Eerkens and Bettinger, 2001; Eerkens and Lipo, 2005) has cogently argued that, in practical terms, when we add limitations in human skill to those of perception, CV values  $\leq 5\%$  may simply be explained by human copying error. An upper limit corresponding to near-random and completely non-standardised production has been modelled, which is a CV of  $\approx 58\%$  (Eerkens and Bettinger, 2001). The range between the two thus provides the scope for varying levels of standardisation. A low CV value represents limited variation and thus standardisation of artefact production, where a template or standard would have been used to achieve the desired outcome.

CV has successfully been applied in previous studies of weaponry to assess standardisation of production along with multivariate methods, by measuring and comparing the variability of key dimensions and features (Li et al., 2014; Martinón-Torres et al., 2014), even if the weapons studied were made of bronze cast in moulds and therefore not comparable to the wrought weapons studied here. Whilst this approach, common in previous studies (see references above) allows for a relative comparison of certain elements of the form, it does not enable a comparison of the overall appearance – the shape. Thus, the study of artefact variability has more often been appreciated in terms of variability of specific dimensions and/or features rather than overall shape. For example, Li et al. (2014) had to supplement metric dimensions with qualitative observations of object shape and microstyle in order to capture a fuller range of variability. This mode of analysis, although insightful, is not the only means of comparing items within a group, whether they were produced by industrial or pre-industrial methods.

## 1.2. Towards GMM analysis: measuring standardisation through shape

It is proposed here that metric analyses may be more reflective of the industrial era, in which precision and accuracy continue to be understood in such terms and considered an *a priori* measurement of standardisation. The previous existence of imperial standards of measurement attests to the variability that once existed prior to globalisation and industrialisation; metric measures have largely become an institutionalised fact worldwide. The modern process of product design and manufacture relies heavily on metric qualities (often cardinal in nature). However, in a pre-modern, pre-industrial era, comparison of objects metrically may not be wholly appropriate. Can we, or should we, judge standardisation of pre-industrial objects in modern terms that come naturally to us? Other means of measuring standardisation should also be considered, explored and studied, especially when assessing the appearance of objects outside the modern 'industrial' context. It is appropriate, therefore, that weapons from the pre-modern era are also investigated in terms of their overall appearance, their shape, and not their metric dimensions alone. One means of achieving this is by GMM analysis, which performs an easier separation between size and shape, allowing for the shape to be investigated independently.

Rather than any given metric characteristic(s), consumers and producers may instead have relied upon (or been motivated by) the overall appearance ('shape') as a benchmark standard. This does not mean to negate size or width as unimportant; rather, it considers that shape could have been the primary function of setting a standard, not the inherent size. 'Shape' is defined as "all the geometrical information that remains when location, scale and rotational effects are filtered out from an object" (Stegmann and Gomez, 2002, p. 2).

Investigations of prehistoric and historic weapons of iron and bronze normally utilise metric measurements to infer their degree of

standardisation (for examples using the Havor lance, see Bemmann and Bemmann, 1998a, 1998b; Ilkjær, 1990a, 1990b; Ørsnes, 1988). This paper presents the first use of GMM analysis as applied to iron weaponry to investigate standardisation in terms of the variation of overall 'shape'.

GMM analysis has become widely used within archaeological science as a means to understanding the variation and relationship between shapes. It has been used extensively within zooarchaeology, evolutionary anthropology, archaeobotany and paleoenvironmental ecology to investigate taxonomic relationships and species identifications (Bignon et al., 2005; Cucchi et al., 2011; Curran, 2012; Evin et al., 2013; Martinón-Torres et al., 2006; Owen et al., 2014; Ros et al., 2014; Seetah et al., 2014; Terral et al., 2004). It has also been useful in studying ceramic typologies (Wilczek et al., 2014) as well as temporal and technological investigations of lithic technologies (Archer and Braun, 2010; de Azevedo et al., 2014; Okumura and Araujo, 2014), with recent developments towards 3-dimensional GMM analysis (Lycett et al., 2010; Lycett and von Cramon-Taubadel, 2013). However, only a few GMM studies exist for metal artefacts (Odler, 2016, p. 248; Wilczek et al., 2015).

Whilst GMM applications have often explored the phylogeny of shapes, looking at their development and evolution, the focus in this paper is to explore the consistency of a single weapon type found at different contemporaneous archaeological sites. The aim is to empirically test its definition as a standardised product, while exploring methodological approaches that may be of use in other studies.

## 2. Material

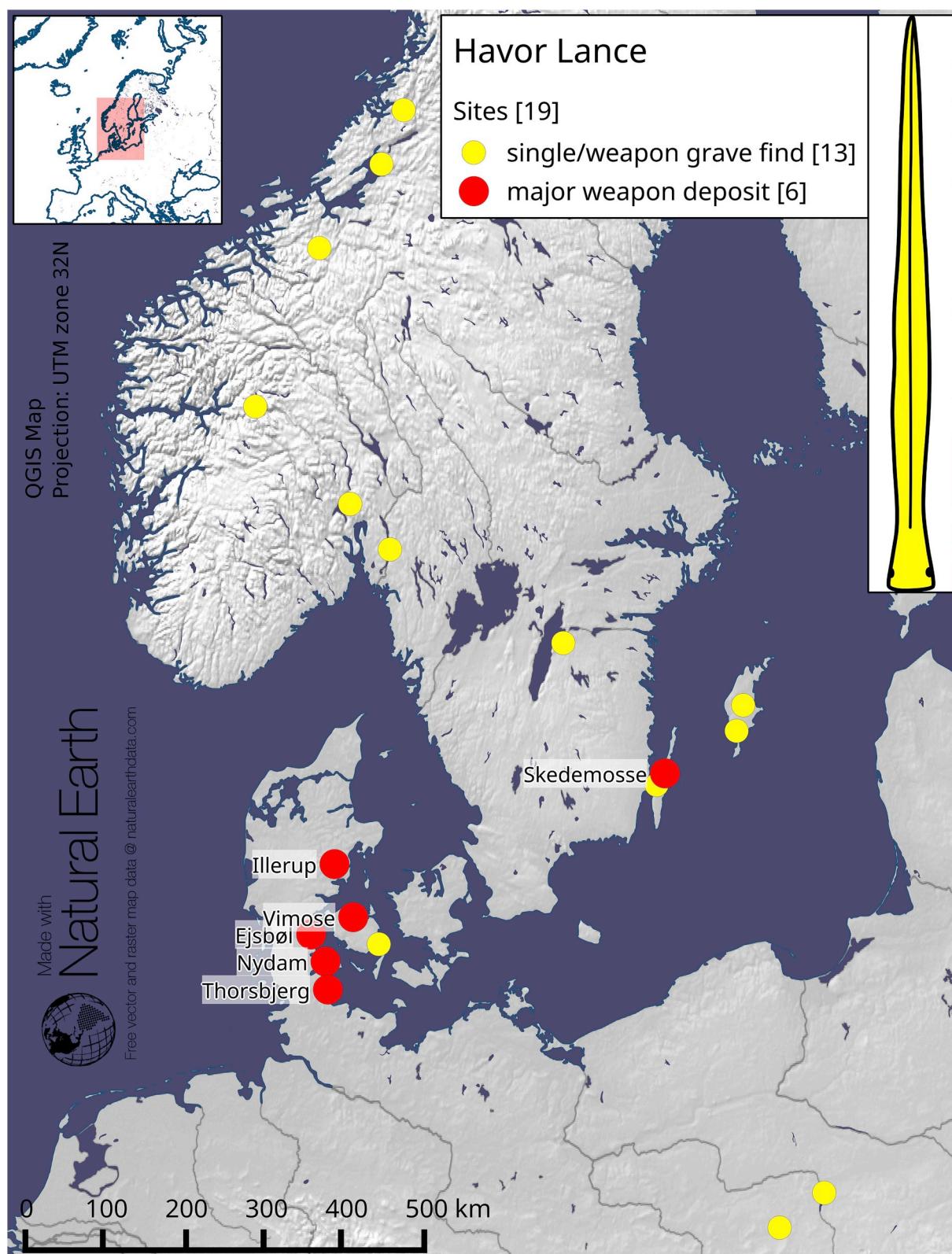
The 'Havor' lance is used here as a case study to illustrate GMM analysis as a useful tool for studying standardised weaponry. This lance type, amongst numerous other standardised weapon types, is found in several substantial weapon offerings (sometimes referred to as war booty sacrifices) known from Southern Scandinavia during the Iron Age. The weapon is more accurately defined in English as a spear head, which would be attached to a wooden staff and thrown as a javelin. However, the term 'lance' (more often used in English to refer to a spear used by mounted horseman), is preserved here, as English publications and translations often refer to this type of spear as a 'lance-head' (cf: Ilkjær, 2000), deriving from the German 'Lanzenspitzen' (cf: Ilkjær, 1990a, 1990b; Ørsnes, 1988) or Danish 'lansespids' (cf: Engelhardt, 1865; Petersen, 1995).

### 2.1. Archaeological background

Nearly 30 weapon deposits are known from across Southern Scandinavia, with the major weapon deposits often containing hundreds, and in some cases thousands of military items. They are thought to represent defeated armies whose weapons were votively deposited (sometimes numerous depositions at a single site over a period of time), mostly dating from around the early third century CE into the Migration Period (c. 500 CE) (Brønsted, 1960, pp. 228–233; Hansen, 2007, 2003; Ilkjær and Lønstrup, 1982; Pauli Jensen, 2010, 2009).

Amongst the weapons, over 120 Havor lances have been found at four of the six major weapon offering sites, from broadly contemporaneous deposits: Ejsbøl, Nydam, Illerup and Skedemosse (Ilkjær, 1990a, p. 53). This study focuses on the examples available to study published from Ejsbøl, Nydam and Illerup, three of the main major weapon deposits along the east coast of Jutland (Denmark). As per the other weapon types found in the deposits, the Havor lance has traditionally been considered a 'standardised' product. However, this assertion has only been investigated for Havor lances from Illerup in the form of histograms (Ilkjær, 1990a).

Whilst most Havor lances are known from these major weapon deposits, other archaeological examples exist as single finds or grave finds from 15 other sites (see Fig. 1), predominantly found outside



**Fig. 1.** Distribution map of Havor lances found as single/grave finds and the locations of the major weapon deposits where the majority have been found, made using QGIS (QGIS Development Team, 2016).

Jutland: six from Norway, six from Sweden, two from Poland and one case from Denmark (Ilkjær, 1990a, p. 58). Ilkjær's (1990a, p. 328) seriation dating of different weapon types, dated by grave finds, along with a relative chronology of other European weapon finds, places the Havor lance in weapon groups 10 and 11, dating it to the second half of

the period C3 (Late Roman Iron Age) and into D1 (Early Germanic Iron Age/Migration Period), corresponding to 375–400 CE.

Many of the personal ornaments (i.e. combs, strike-a-lights, brooches) found as part of the *militaria* in the weapon deposits are distinctly Scandinavian in origin (Hansen, 2003) and even some of the weapons

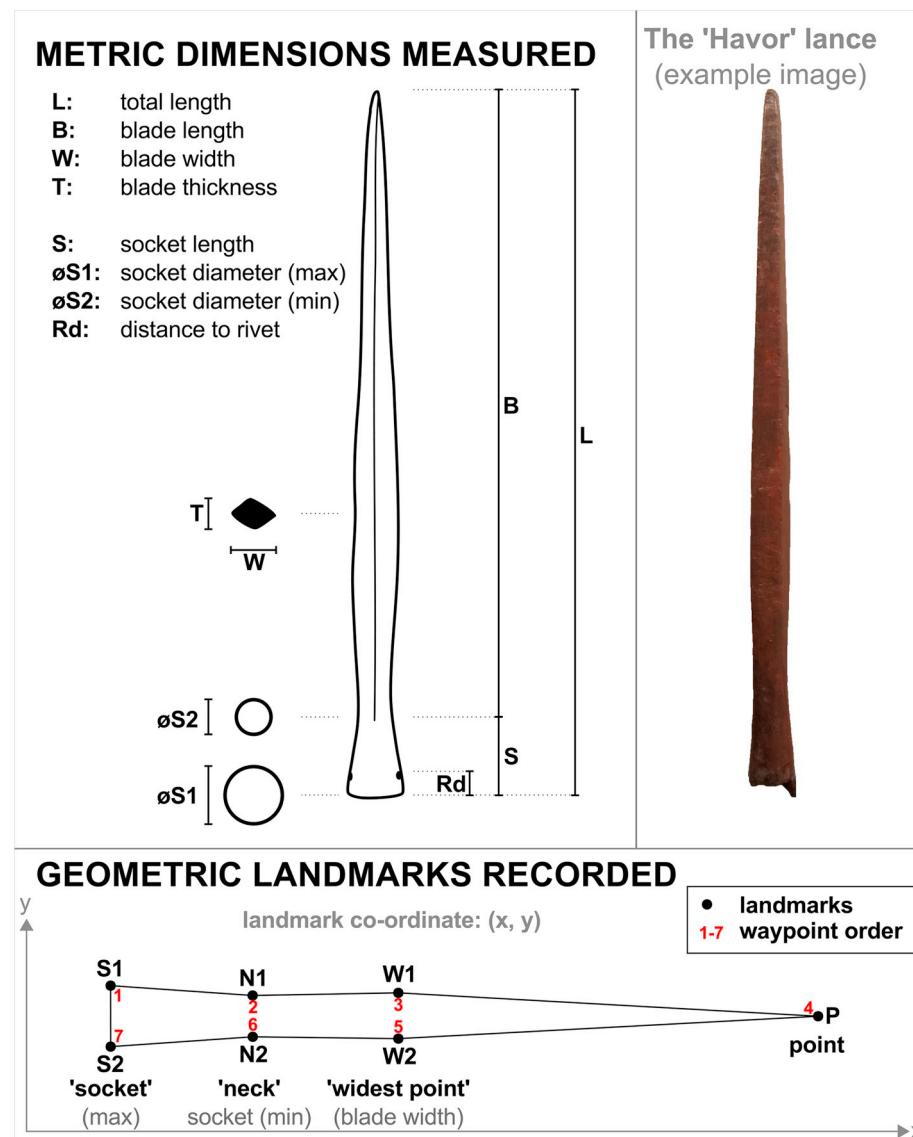


Fig. 2. Illustration of the Havor lance form and features highlighting the metric dimensions and geometric landmarks recorded.

themselves are considered 'Nordic types' (Hansen, 2003, p. 87; Ilkjær and Lønstrup, 1982, p. 63), providing an indication as to the origins of finds.

## 2.2. Smithing technique and iron sources

As part of the broader research into the Havor lance, 13 specimens were available for metallographic analysis from two of the main sites (Nydam and Illerup), where their metal was chemically analysed along with their entrapped slag inclusions, and the microstructure studied. The results (Birch, 2018), reveal that the 13 lances studied were produced out of bloomery iron, hammered using the same spiral-form construction technique, however the metal (either phosphoric iron or steel) is representative of different sources of iron being utilised (conforming to three broad origin groups) likely deriving from Norway and Jutland itself. Thus, whilst the Havor lance appears outwardly standardised, and so does the smithing technique used for construction, different sources of iron were being used as the material for manufacture.

## 2.3. Weapon workshops?

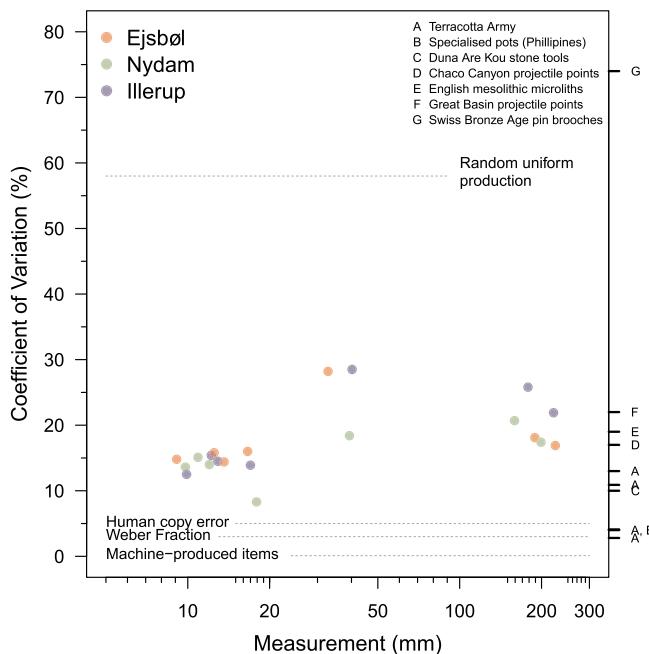
Several of the Roman-manufactured swords found in the weapon deposits bear Latin inscriptions that have been interpreted as manufacturers' marks, and it has been suggested that the runic inscriptions found on spearheads are an imitation of this practice, to denote quality (Imer, 2010, pp. 53, 55). Three lances from Illerup of the same (Vennolum) type bear the exact same runic inscription 'wagnijo', which has been interpreted as a manufacturers' mark, either that of the weapon smith, the workshop or owner of the workshop, and the runic inscription supports the interpretation that they were produced in the Scandinavian area (Imer, 2010, pp. 54–55). It has also been suggested that 'Wagnijo' may have been one of the leaders of the attacking forces (Ilkjær, 2000, p. 116), so one might interpret these inscriptions as ownership marks of war gear being dispensed from a centralised authority.

Two other different runic inscriptions have been found on two other lance heads also interpreted as manufacturers' marks (Imer, 2010, pp. 56–61). Although none of the Havor lances bear any manufacturers' marks or inscriptions, their tight chronological span could be consistent with a single or small number of specialised workshops.

**Table 1**

Summary data of lance metric measurements (in mm) and coefficients of variation (in %), showing the number of samples analysed (n = ).

Dimension	Ejsbøl	Nydam	Illerup	Overall
Total length	Mean (mm)	224.8	± 38.0	221.4
	CV (%)	16.9		21.9
	n =	6		24
Blade length	Mean (mm)	189.0	± 34.2	178.3
	CV (%)	18.1		25.8
	n =	6		30
Blade width	Mean (mm)	13.6	± 2.0	12.9
	CV (%)	14.4		14.5
	n =	16		48
Blade thickness	Mean (mm)	9.1	± 1.3	9.9
	CV (%)	14.8		12.5
	n =	16		43
Socket length	Mean (mm)	32.8	± 9.2	40.2
	CV (%)	28.2		28.5
	n =	13		31
Socket thickness (max)	Mean (mm)	16.6	± 2.7	17.0
	CV (%)	16.0		13.9
	n =	12		44
Socket thickness (min)	Mean (mm)	12 ≈ .5	± 2.0	12.2
	CV (%)	15.8		15.4
	n =	15		42



**Fig. 3.** Havor lance standardisation (CV%) shown in relation to other standardised/non-standardised archaeological examples, marking the range (Weber Fraction-Random production) of standardisation.

### 3. Methods

A total of 123 lances from Ejsbøl (n = 17), Nydam (n = 49) and Illerup (n = 57) were recorded using existing tabulated data or scaled figure plates (drawings and photographs) from published excavation reports (Bemmam and Bemmam, 1998a, pp. 180–181, 1998b, pp. 102–106, 1998b, Figs. 103–107, Ilkjær, 1990a, pp. 53–59, 1990b, Figs. 13–21, Ørsnes, 1988, pp. 65–71, 1988, Figs. 128 and 136). We acknowledge that imprecisions caused by the multiple hands and recording methods behind our data may have added some noise in our dataset, with the likely effect that variability as determined here is likely higher than real. At the same time, the use of published data was deemed the most economic approach, and one that could allow us to test the use of comparable data in other contexts. As an agenda for the future, the extent of this analytical error should be quantified e.g. by

comparing the results on a sample of objects, each recorded by a variety of means.

Eight specific measurements were recorded on each specimen, corresponding to the overall dimensions and distance of rivet-hafting of the Havor lance. Although socket 'diameters' were measured, they are referred to here as minimum and maximum socket 'thickness', as not all lance sockets were circular in shape. Rivet distance could only be measured in 55 lances; it was not possible to discern whether this feature was present on all lances due to the condition of the sockets or angle of recording visible from plate figures and drawings. The values showed a very wide spread but we cannot be confident about the reliability of this data. Hence we report this data in Appendix 2 but we will not discuss this feature further.

GMM analysis was conducted on 78 of lances considered 'complete' from Ejsbøl (n = 6), Nydam (n = 24) and Illerup (n = 48), where all features could be observed. Seven landmarks were selected for defining the Havor lance shape and for performing GMM analysis.. Landmarks are universal points "of correspondence on each object... matches between and within populations" (Stegmann and Gomez, 2002, p. 2). In order to compare lance shapes, all co-ordinates were superimposed using a General Procrustes Analysis (GPA); coordinates of all weapons are first translated (so that all share the same gravity center), scaled by size and then rotated in order to minimise the distance between homologous landmarks (Goodall, 1991; Gower, 1975). See Fig. 2 for the metric measurements and geometric landmarks recorded in this study.

Lances were first digitised in 2 dimensions from drawings or photographs using 'geomorph 2.1.1' (Adams and Otárola-Castillo, 2013). Subsequent GMM analysis was performed using the dedicated library packages 'geomorph 3.0.6' and 'shapes 1.2.3' (Dryden, 2013) in the R environment (R Core Team, 2017) for statistical computing (version R 3.4.0 2017-04-21). A full annotated R script of the analyses can be found in Appendix A, all traditional metric data in Appendix B and all GMM co-ordinates in Appendix C.

Metric data was investigated using descriptive methods. The CV was calculated for each measurement recorded for the lance overall and by site, followed by an examination of Pearson's product-moment correlation coefficients (Pearson's *r*) between lance features. In this paper Pearson's *r* is classified as being weak ( $0.1 \leq r \leq 0.5$ ), strong ( $0.5 \leq r \leq 0.9$ ) or very strong ( $0.9 \leq r \leq 1$ ).

GMM analysis started with a GPA for visual inspection and mean shape difference tested statistically using the Hotelling  $T^2$  statistic, the Goodall F test statistic and a Procrustes Analysis of Variance (ANOVA),

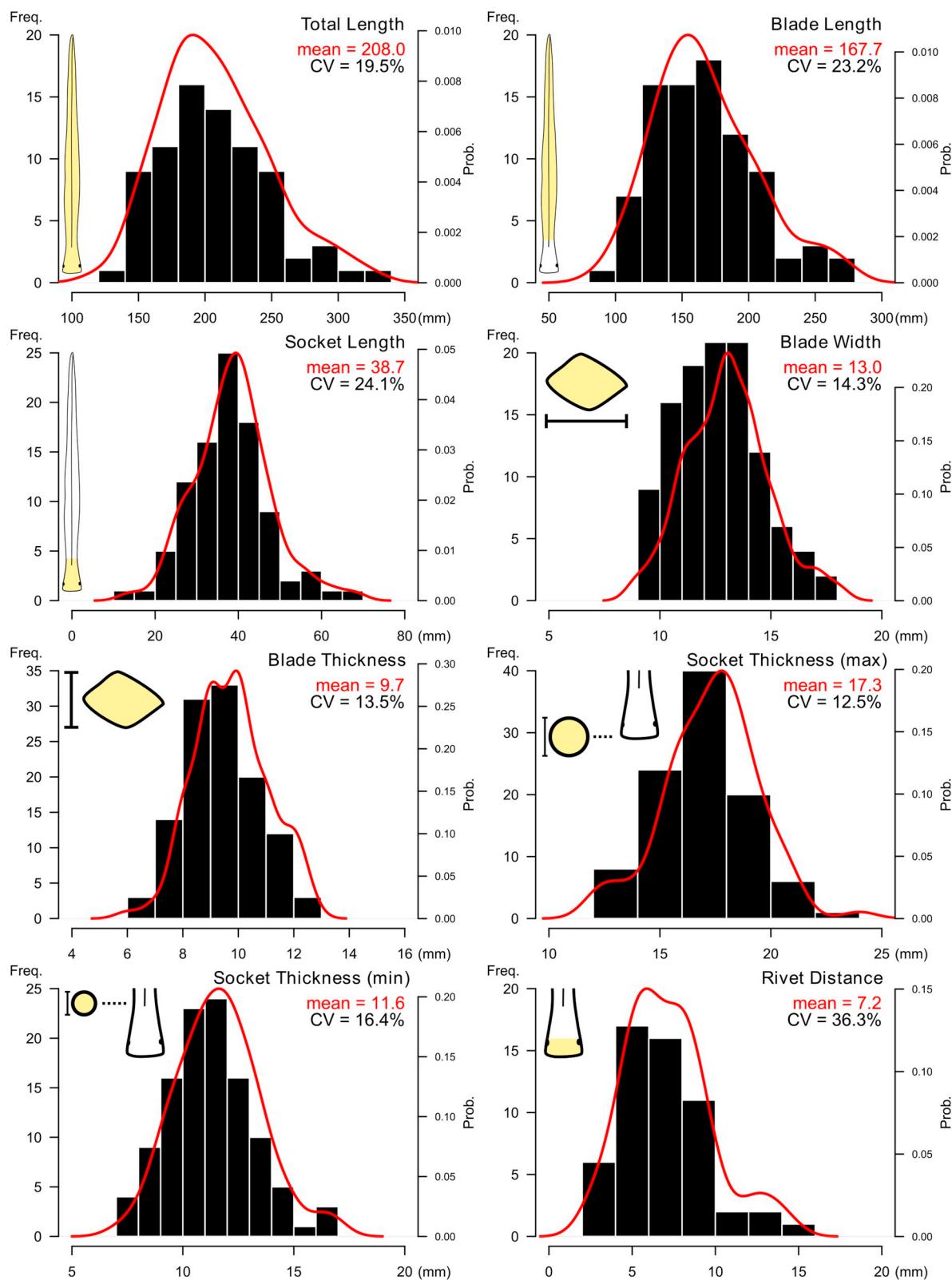
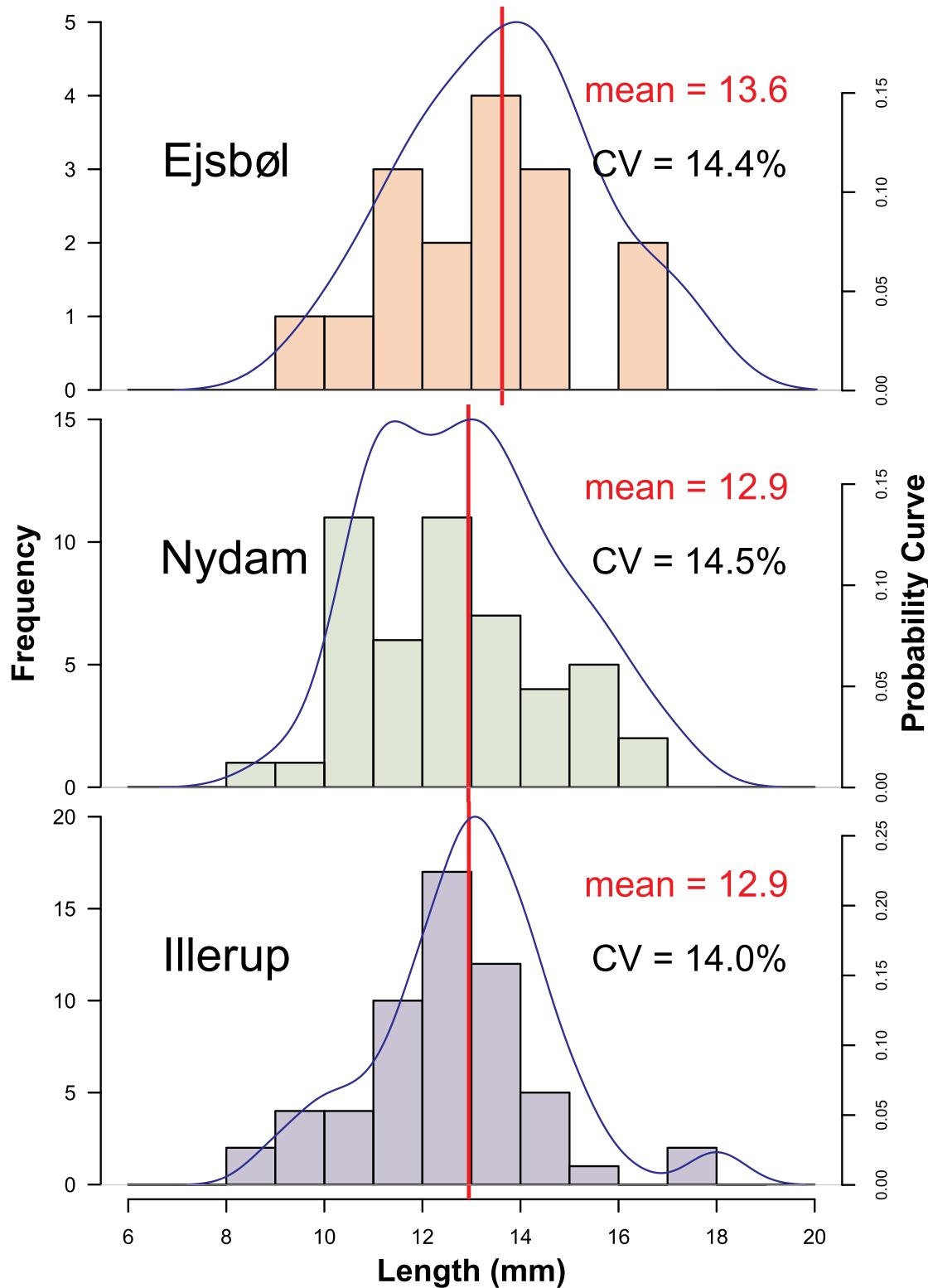


Fig. 4. Histograms of the 8 Havor lance dimensions recorded on 123 lances, showing the mean (mm), CV (%) and kernel density estimation (Prob.).

which is equivalent to a permutational-MANOVA (Multivariate Analysis of Variance).

The Hotelling  $T^2$  statistic is used to test the (multivariate) means of different populations, in what is essentially a multivariate form of a (univariate) Student's  $t$ -test for testing two samples. Any p-value obtained lower than our significance level ( $\alpha = 0.05$ ) would force the null

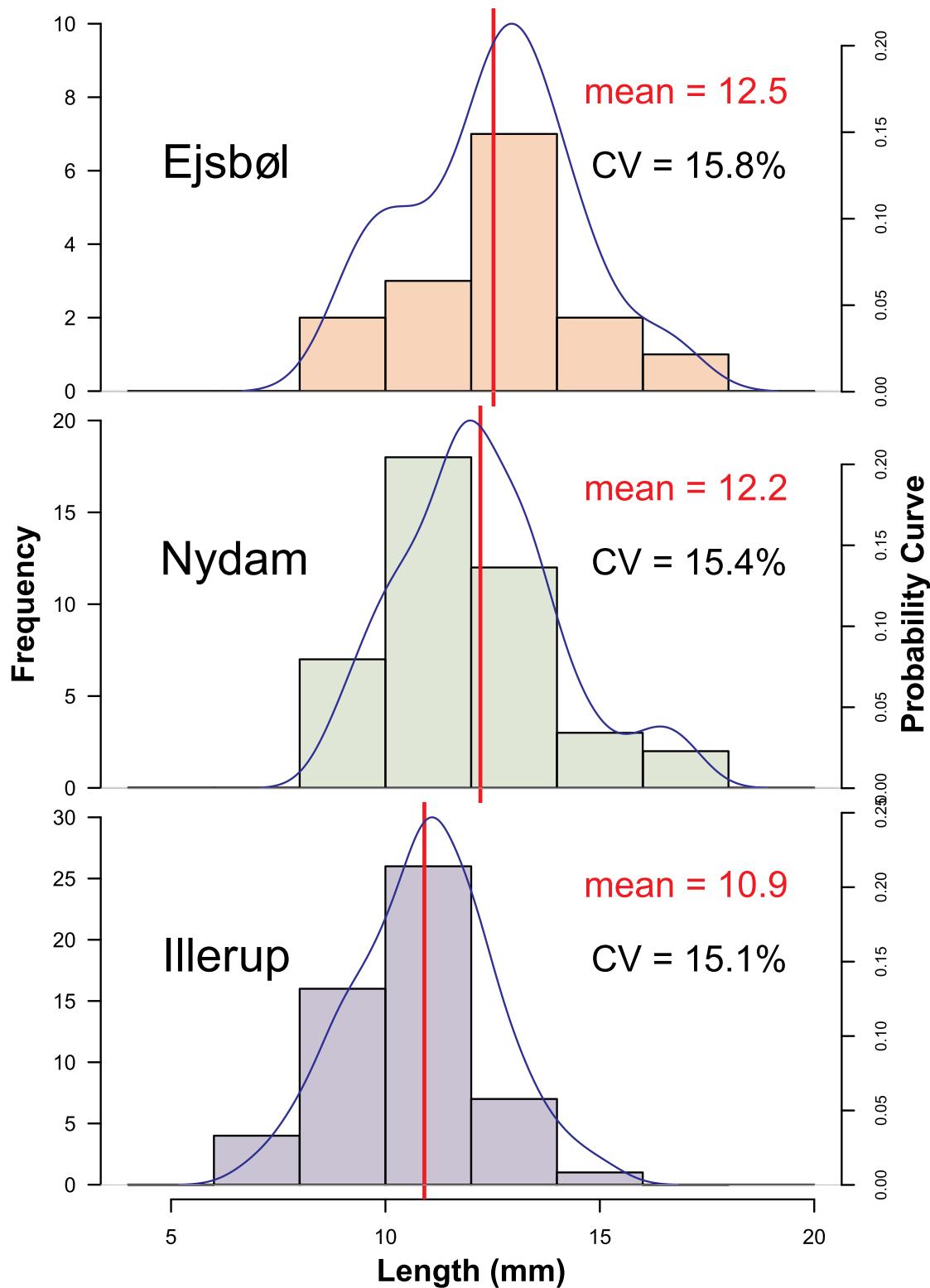
hypothesis (no significant difference in lance shape) to be rejected. As only two samples are compared, the Hotelling  $T^2$  statistic is obtained for the pair-wise combinations between the three sites studied. The Goodall F test is used to test mean shape difference between two samples based on the variation within each sample. This is different to the Hotelling  $T^2$  statistic, which is based on multivariate means. In order to determine



**Fig. 5.** Histogram of blade width measurements recorded from each site, showing the mean (mm), CV (%) and kernel density estimation (Prob.).

whether there is a real difference in mean shape based on the factor variable 'sites', a Procrustes ANOVA was performed on the GPA of the 77 Havor lances. This allows for all three groups to be compared simultaneously based on their variance, providing a more powerful and apt means of testing mean shape. The null hypothesis is that all group means (shape) are equal.

The F statistic reported via these techniques is presented in the following format:  $F[df,t] = [\text{result}]$  where  $df = \text{degrees of freedom}$  and  $t = \text{total number analysed}$  and result = F statistic obtained. The p-values reported refer to the significance level obtained. The null hypothesis is that there is no significant difference in mean shape between Ejsbøl, Nydam and Illerup ( $\alpha = 0.05$ ). GMM data was subjected to

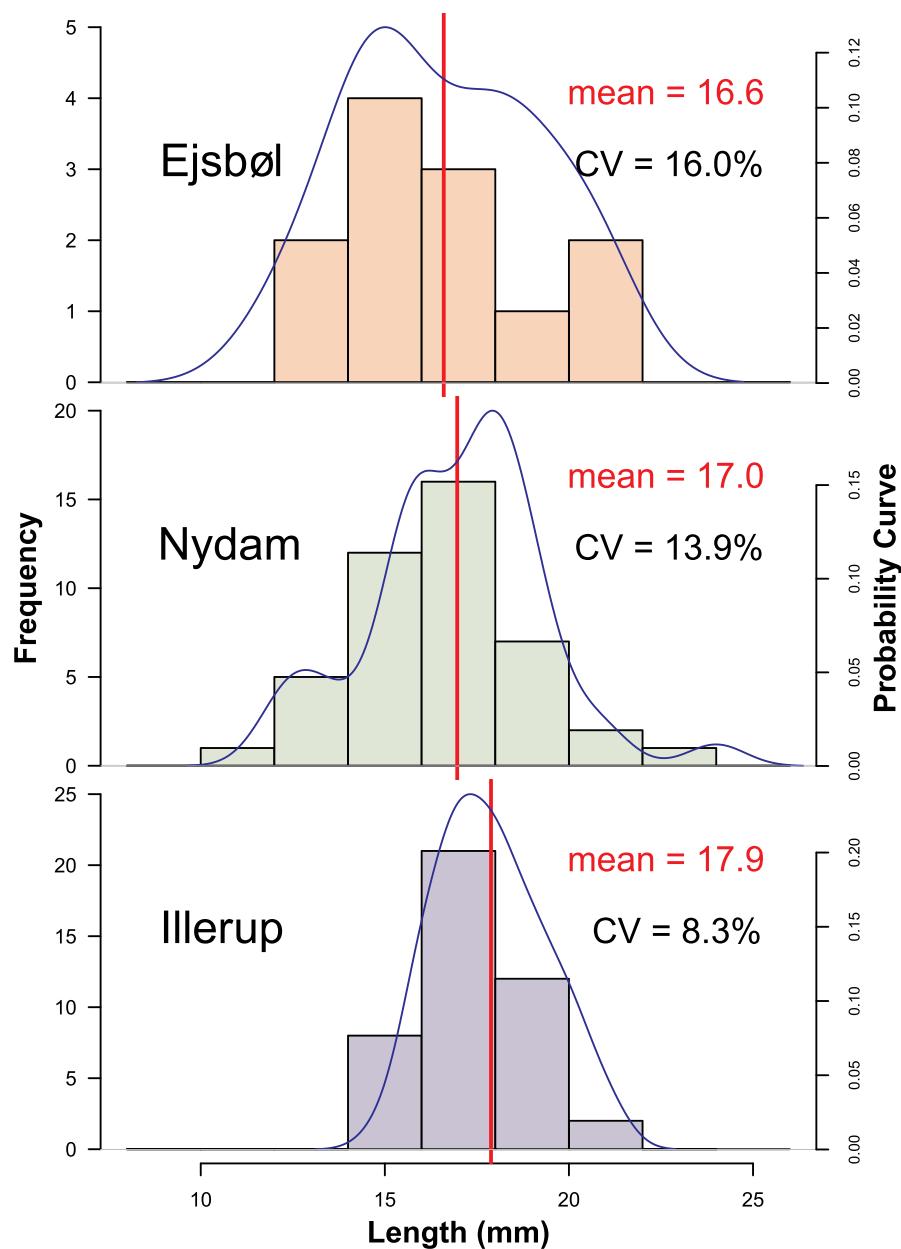


**Fig. 6.** Histogram of minimum socket thickness measurements recorded from each site, showing the mean (mm), CV (%) and kernel density estimation (Prob.).

further multivariate analyses and investigation of object symmetry, which was performed using 'geomorph 3.0.6'. The multivariate methods performed on both traditional metric and GMM data were hierarchical cluster analysis (HCA) and principle components analysis (PCA). The robust statistical approach adopted here justifies using multiple statistical methods to assess broadly for similar outcomes.

#### 4. Results: metric analysis

The CVs calculated of the 8 measurements recorded from 123 Havor lances are indicative of a relatively high degree of dimensional standardisation (Table 1). CVs are relatively lower for 'width' measurements such as blade width, blade thickness and socket thickness



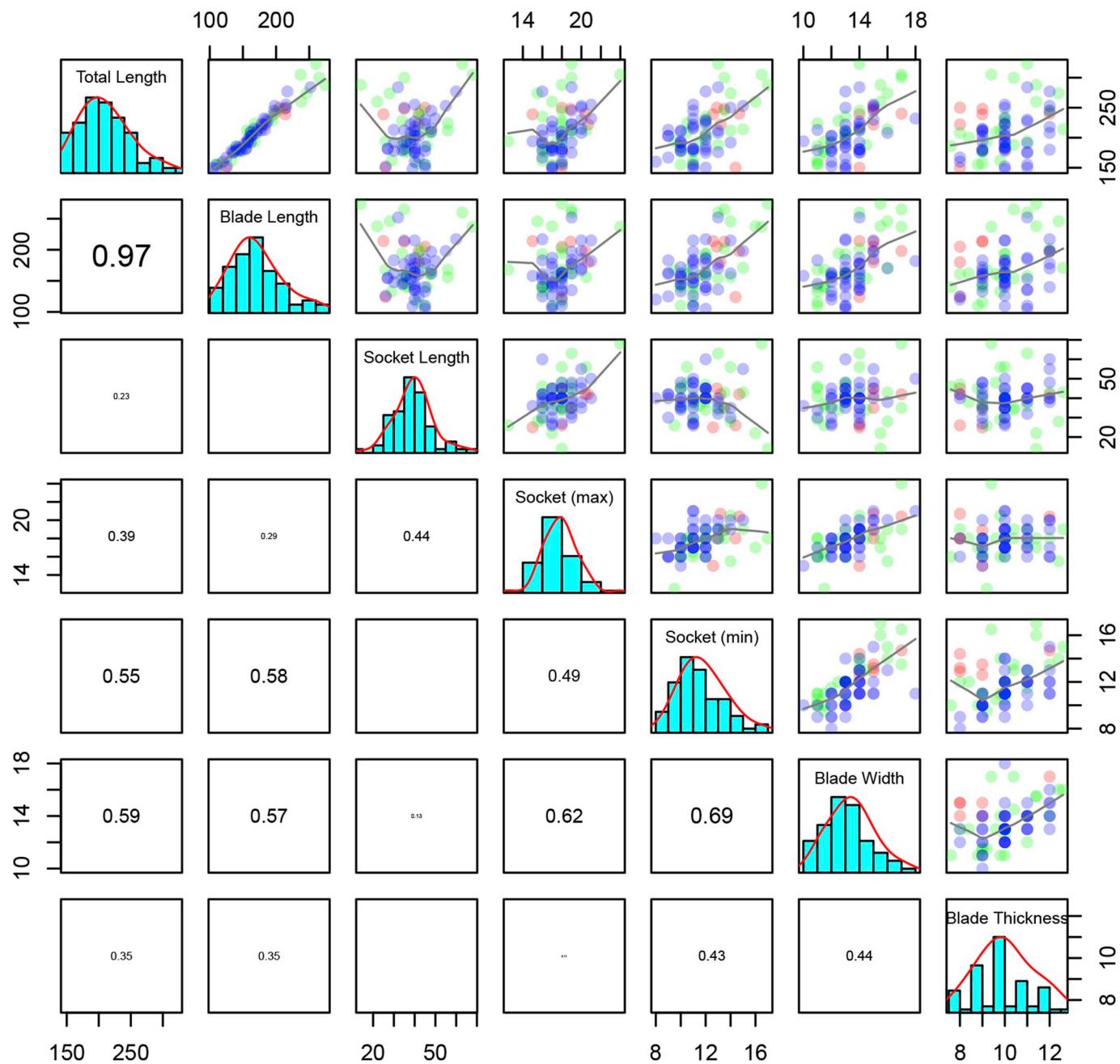
**Fig. 7.** Histogram of maximum socket thickness measurements recorded from each site, showing the mean (mm), CV (%) and kernel density estimation (Prob.).

( $CV \approx 14\text{--}16\%$ ) compared to length ( $CV \approx 20\text{--}24\%$ ). This could indicate that uniform length was either less important or harder to achieve, although it is possible that the higher variability may at least in part be due to larger errors in recording and measurement by modern researchers. Fig. 3 shows that greater CV values are generally associated with larger dimensions. It also shows the relatively highly standardised nature of the Havor lance, by comparing CV values to those recorded in other archaeological case studies (Fig. 3).

Not visible in Table 1 are the distributions of all measurements recorded, presented in Fig. 4 in the form of histograms with probability curves (kernel density estimation). Although the histograms reveal a largely normal distribution and confirm the Havor lance to be largely standardised in dimensions and likely imitating a single model, some of the lengthwise proportions (total length, blade length, socket length) show outliers, which are discussed in more detail later. The uniformity in dimensions of the Havor lance is further emphasised by comparing sites directly, especially in the distribution of width proportions, such as

the blade width (Fig. 5), minimum socket thickness (Fig. 6) and maximum socket thickness (Fig. 7). Not only are the normal distributions overlapping and means comparable, but the low CVs reported (14.3%, 13.5% and 12.2% respectively) correspond to differences in the range of 1–2 mm in real terms.

As per previous studies, investigating CV only allows for a discussion of certain features or dimensions in isolation. This can be improved by studying two-way pair combinations of measurements in the form of a correlation matrix, shown in Fig. 8, where Pearson's  $r$  values are accompanied by bivariate scatterplots of raw data and histograms of the distributions (with kernel density estimate curves). Pearson's  $r$  are strong for width dimensions (blade width and both socket thicknesses), whilst the weakest coefficients were obtained for the socket region of the lance. The correlation matrix clearly demonstrates many of the lance features to be related, with an obvious linear relationship between blade length and total length (linear regression score = 0.94).



**Fig. 8.** Correlation matrix of Havor lance dimensions showing Pearson's  $r$  (lower panel), accompanying bivariate plots (upper panel) and distributions (histograms with kernel density estimation, on the diagonal). All pairwise observations were analysed from the original dataset. Red = Ejsbøl, green = Nydam, blue = Illerup. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## 5. Results: GMM analysis

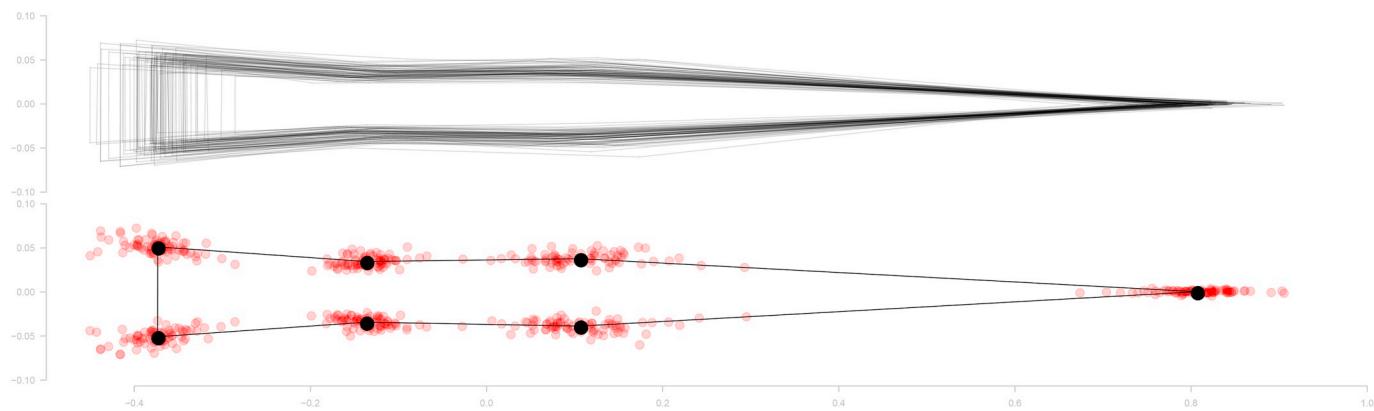
The GPA of all complete lance head shapes and landmarks recorded (superimposed) are shown in relation to the overall mean shape in Fig. 9, and the GPA of lances for each site can be seen in Fig. 10. The impression of uniformity in Havor lance shape can be visualised, as individual lances show great conformity to the mean shape. The two main differences are in the socket region and the point, where variation in the relative position of landmarks can be observed. This corresponds with the x co-ordinate value of landmarks, being the length-axis of the lance shape. Conversely, landmarks show very little deviation along the width axis (y co-ordinate value), showing good agreement with the results from metric analysis.

When sites are compared directly, the GPA superimposition of

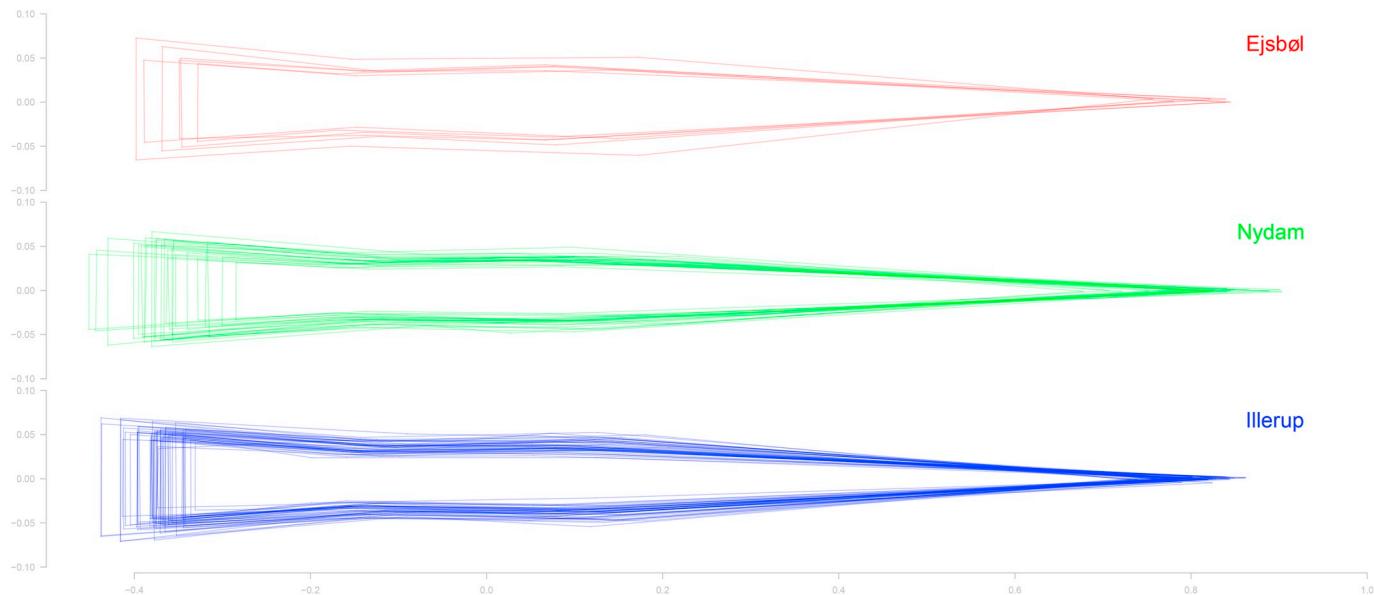
Havor lance shapes allows for a clearer identification of areas of disparity, notably the socket area and elongation of the point (pronounced in the case of Nydam). The side-by-side comparison also contrasts the distinct lack of curvature in some of the Nydam lances (potential outliers) with the 'classic' Havor lance shape, which shows the re-curve at the 'neck', clearly visible in Ejsbøl and Illerup. The differences apparently displayed in the Nydam assemblage correspond with a discrete population of lances from this site with larger metric values, confirmed in the following sections.

### 5.1. Testing mean shape: Hotelling $T^2$ statistic, Goodall F test and Procrustes ANOVA

The differences observed in Havor lance shape thus far are purely



**Fig. 9.** GPA superimposition of complete Havor lances illustrated with the mean shape estimate, plotted alongside all landmarks recorded.



**Fig. 10.** GPA superimposition of complete Havor lances from each site.

descriptive and must be accompanied by a statistical comparison of mean shape difference. Prior to doing so, it is necessary to establish whether there is a significant relationship between centroid size and shape, by monitoring the allometric shape variation. There was found to be an association between overall centroid size and shape ( $F_{1,77} = 5.3$ ,  $p = 0.015$ ), however this could not be explained through site difference ( $F_{2,77} = 0.3$ ,  $p = 0.080$ ) (see the R script for further information).

Two of the three Hotelling  $T^2$  test results proved not significant. The mean shape difference between Ejsbøl and Nydam ( $F_{2,27} = 3.6$ ,  $p = 0.099$ ) is insignificant, and so is the difference between Illerup and Nydam ( $F_{2,69} = 0.7$ ,  $p = 0.739$ ). The Hotelling  $T^2$  statistic was significant for Ejsbøl and Illerup ( $F_{2,49} = 5.7$ ,  $p < 0.004$ ). As only one of the three site combinations yielded a significant difference in multivariate means, a second different statistical test was conducted.

The Goodall F test confirmed once more that there is no significant difference in shape between Ejsbøl and Nydam lances ( $F_{2,27} = 0.084$ ,  $p = 0.806$ ), as well as Illerup and Nydam lances ( $F_{2,69} = 0.829$ ,  $p = 0.490$ ). Unlike the Hotelling  $T^2$  statistic, the Goodall F test also confirmed that there is no significant difference in mean shape between Ejsbøl and Illerup ( $F_{2,49} = 0.752$ ,  $p = 0.430$ ), indicating that the small sample size of Ejsbøl may be responsible for the differing results.

Two-way significant testing for mean shape difference has confirmed there to be no significant difference (in terms of multivariate means and internal variation) between Ejsbøl and Nydam, and Illerup

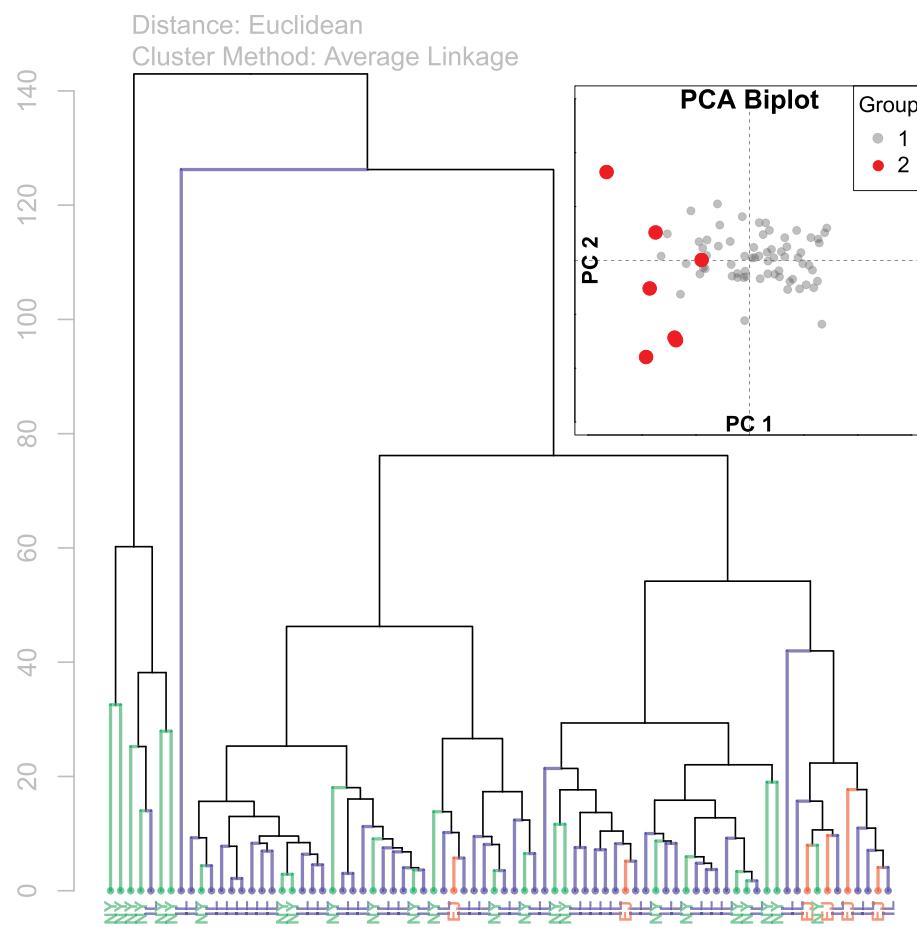
and Nydam, with conflicting results for the Ejsbøl/Illerup comparison. The result of the Procrustes ANOVA was insignificant ( $F_{2,77} = 0.592$ ,  $p = 0.7$ ). Therefore, we accept our null hypothesis that there is no significant difference in mean shape between Ejsbøl, Nydam and Illerup.

## 6. Multivariate analysis: a comparison of metric and geometric morphometric data

Multivariate statistical analysis was conducted on both metric and GMM (GPA aligned co-ordinates) datasets. The results are compared here, with the aim of evaluating the usefulness of each data type as a means for investigating variation (and thus standardisation) using multivariate methods. The multivariate investigation also permits a visualisation of any variation in a reduced number of dimensions.

### 6.1. Hierarchical cluster analysis (HCA)

HCA of both datasets provides strikingly similar results, with the metric data presented in Fig. 11, and the GMM data shown in Fig. 12 (the different HCA scales arise from the different datasets being used). Both HCA dendograms reveal two main groups. The first cluster, defined by a distinctively high breakage in each dendrogram, consists of six lances from Nydam (though Fig. 11 also includes one lance from Illerup). These six lances are identified as 'Group 2' in the embedded



**Fig. 11.** HCA dendrogram of lance metric measurements, with embedded PCA (see Fig. 13).

PCA diagrams shown in Figs. 11 and 12, which are miniaturised versions of the proceeding Figs. 13 and 14.

The breakages for the second and third branches in both HCA dendograms are not much lower, indicating that any difference to the first cluster is slight. None of the remaining clusters correspond with site, indicating that lances are indistinguishable (in terms of site affiliation) based on their metric dimensions or shape. The only difference that can be found is in the sub-group of six Nydam lances identified.

## 6.2. Principal component analysis (PCA)

The quality of the PCA results obtained from using GMM (GPA aligned) co-ordinates is a significant improvement on the PCA results obtained by metric data alone. The PCA biplot of metric data (Fig. 13) of the first two principal components explains 70% of the total variation, which might be considered unsatisfactory by many statisticians. The PCA of metric data shows no correspondence between lances and any particular feature(s) or dimension(s) based on site grouping. It could be argued that the total variation explained by the first two components is low due to insufficient differences between Havor lances from different sites; it indirectly attests to the highly standardised nature of the Havor lance. The six extreme cases observed in Fig. 13 correspond to the sub-group of Nydam lances identified in the previous HCA.

The PCA of GMM data (Fig. 14), however, explains an impressive 98% of the total variation, marking for a great improvement on the previous PCA of metric data. The loading vectors scatter evenly around zero. The seven landmarks used equate to 14 variables (where the x and y co-ordinate values are separated). The PCA clearly confirms that the

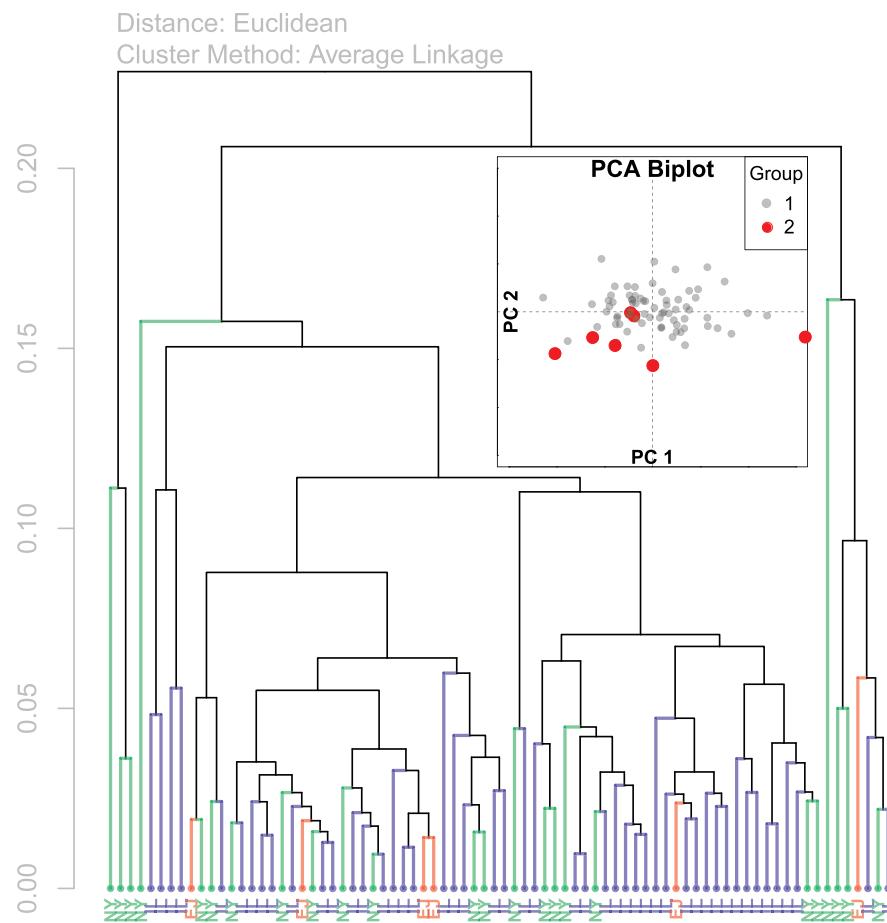
variation is explained by the x co-ordinate values of the point (Px), socket (S1x, S2x) and width (W1x, W2x), which all refer to the slenderness of the Havor lance shape. Almost none of the observed variation is explained by the width-axis ( $PC_2 = 6.3\%$ ). As per the previous PCA, the six most extreme cases shown along PC1 all correspond to the subgroup of six Nydam lances previously identified.

Fig. 14 is accompanied by two warpgrids. The two warpgrids reference the two most extreme cases (negative and positive) on PC1 against the mean shape, resulting in a distortion (or 'warping') of the relative landmark positions. The warpgrids serve to emphasise visually the shape variation observed, which in this case is due to the length-wise landmark positions (x co-ordinate values), the slenderness of the shape. The warpgrids show no distortion along the width-axis of the extreme cases (against mean shape).

The results from the PCA analyses highlight that GMM data is more useful than metric data for discriminating differences. The six Nydam outliers more or less cluster together in the PCA of metric data (Fig. 13), however in the PCA of GMM data (Fig. 14) they form two separate groups reflecting the most extreme shape differences (slender versus chubby) as illustrated by the warpgrids.

## 7. Directional and fluctuating asymmetry of the Havor lance shape

Variation in symmetry of the Havor lance shape is further explored here as an additional means of assessing standardisation. It is assumed that this parameter may have been key to the performance of the weapons. The Havor lance exhibits bilateral symmetry, where both sides of the lance are mirror images of each other, or reflections, with the axis of symmetry bisecting the median plane shape longitudinally



**Fig. 12.** HCA dendrogram of lance shapes (GPA aligned co-ordinates), with embedded PCA (see Fig. 14).

(from point to socket). This is known more specifically as 'object symmetry', where the mirror line runs through the structure of the object (Klingenberg, 2015, p. 848); for a more detailed overview of symmetry and measures of asymmetry that go beyond the scope of this paper, see Klingenberg (2015).

For object symmetry, it is important to be aware of the potential measurement errors when collecting two-dimensional data from three-dimensional objects, where camera alignment or perspective may influence results. It is also important to appreciate that only shape (not size) is investigated for object symmetry, as both sides are inter-dependent and differences are an "aspect of the shape of the whole structure" (Klingenberg, 2015, p. 849).

Asymmetry represents the deviation from symmetry, and so object symmetry is investigated here via directional asymmetry (DA) and fluctuating asymmetry (FA). "Directional asymmetry can be quantified by the difference between the left and right averages", where there is a "tendency" or "systematic difference" between the left and right sides (Klingenberg, 2015, pp. 851–852). FA represents the small differences between an individual and the 'target phenotype', the random or residual differences between left and right sides (Klingenberg, 2015, p. 852). Therefore the differences between the left and right sides provide a measure of asymmetry; for pure fluctuating asymmetry, left and right mirror images should follow a normal distribution with a mean difference of zero (Klingenberg, 2015, p. 850). It is possible that any asymmetry discerned may be caused by taphonomic processes, however it is not possible to investigate this issue further.

As no significant difference was found in Havor lance shape between archaeological sites, all lance shapes were analysed collectively to assess the bilateral symmetry of the overall Havor lance shape. A GPA with bilateral symmetry analysis (BSA) was conducted using the

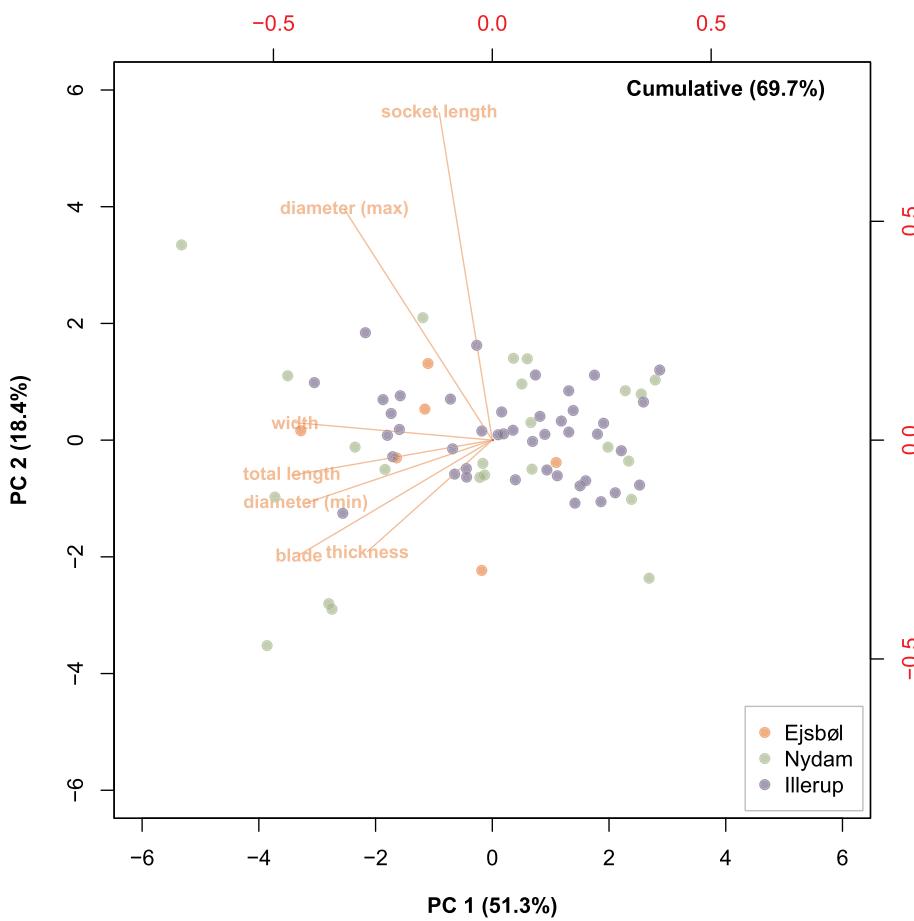
dedicated 'bilat.symmetry' function in the library package geomorph (version 3.0.1); the function performs an analysis of DA and FA for shapes with bilateral symmetry. As stated in the package manual (date 2016-07-18) for geomorph for the bilat.symmetry function, "the shape variation is decomposed into variation among individuals, variation among sides (directional asymmetry) and variation due to an individual  $\times$  side interaction (fluctuating symmetry). These components are then statistically evaluated using a Procrustes ANOVA."

The GPA ANOVA was performed to analyse the variability in DA and FA for Havor lance shape and the results are shown in Table 2. Although significant variation was found between individuals symmetric components ( $p < 0.05$ ), no significant difference was found between left and right sides of the Havor lance ( $p > 0.05$ ), as is visually exhibited by the lack of deformation in the warp grids for DA and FA shown in Fig. 15.

The asymmetric shape data was further processed to obtain a scalar measure of symmetrical difference, which could be used to calculate a CV for symmetry and thus allow comparative studies of symmetry between different objects. For each of the three landmark pairs (landmarks 1 and 7, 2 and 6, 3 and 5), the asymmetrical co-ordinates were compared and their relative differences calculated. As each co-ordinate comprises of an x and y value, two results are obtained for each pair compared. The following calculation was used to compare the difference between x values for a landmark pair, subsequently repeated to calculate the difference between y values:

$$\text{Relative difference} = \left( \frac{a - b}{(a + b)/2} \right) \times 100$$

Where  $a$  represents the value of the first co-ordinate and  $b$  the value of the second co-ordinate. The difference between  $a$  and  $b$  is essentially



**Fig. 13.** PCA biplot of lance metric measurements.

expressed as a fraction of the mean co-ordinate (the mean of  $a + b$ ). The result is the relative difference (%) between co-ordinate values. The differences in x and y values for each of the three co-ordinate pairs are given in Table 3. The median plane bisects the Havor lance through the point (landmark 4) longitudinally, passing through down to the socket (between landmarks 1 and 7). As such, it is the difference between y values that are of most interest, representing the difference between both reflections.

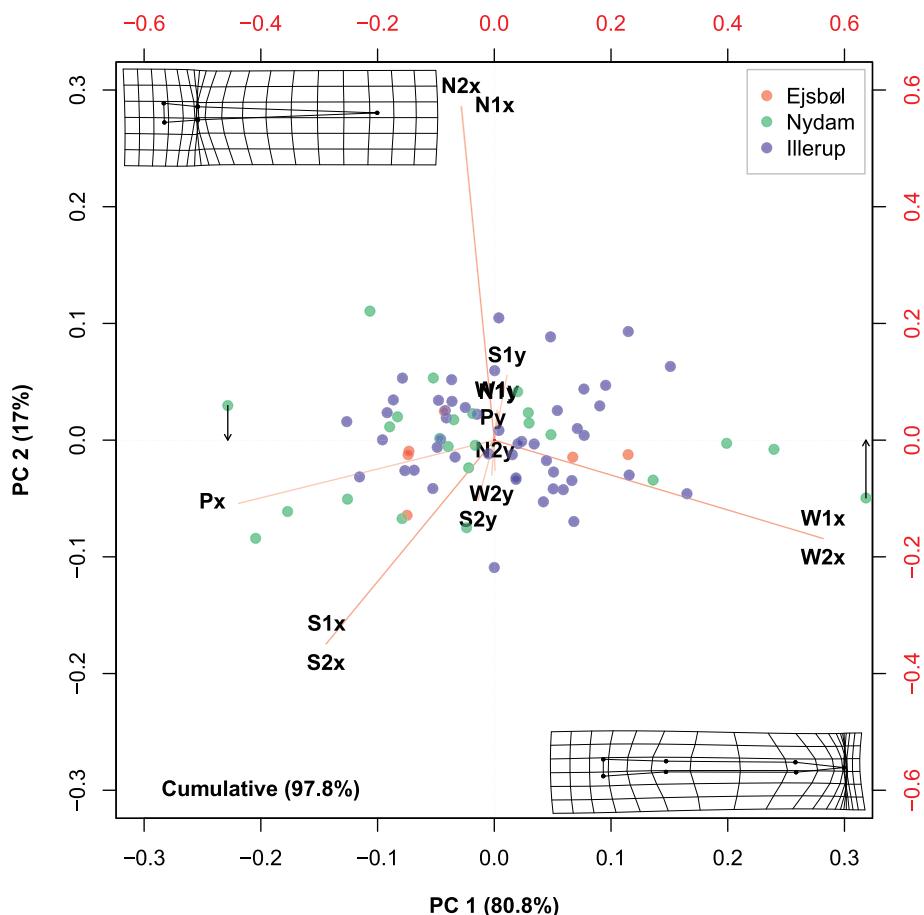
The difference in x values for each of the three pairs is negligible, as this simply represents the position of the point being measured in the reflection. The standard deviation of relative difference in y values for the socket and neck are around between 4 and 5%, whilst these differences are slightly higher for the width, at 7%. In real terms, this translates to a maximum relative difference of 1 mm for the average width (13 mm). These small differences may be attributed largely to copy-errors, and denote the high skill of the smiths who made the Havor lances as symmetrical objects. Overall, the asymmetric differences shown here, expressed as relative difference in co-ordinate values for landmark pairs, may be a useful tool for providing a scalar measure of symmetry.

## 8. Discussion

The results presented in this paper confirm that shape can and should be used as a new avenue by which to investigate artefact standardisation. In the following paragraphs we outline some of the implications of our findings for the understanding of the Havor lance, and explore more generally the potentials and limitations of our approach.

### 8.1. The 'Havor' lance: a standardised product

This study has provided an empirical basis for our understanding of the Havor lance as a standardised product, both in dimensions and in shape. By and large, there are no significant differences within or between sites, confirming the inference of the lance as a highly standardised weapon. Considering that all of these weapons are assigned to a chronological range spanning 25 years only (375–400 CE), and the results of metallographic analyses indicating similar construction methods (Birch, 2018), it may be suggested that a single workshop, or a small number of them, produced all or the majority of the Havor lances recovered. It is interesting to note, however, that the metallographic study points to a range of different metal sources, potentially highlighting meaningful differences between the geological and manufacturing provenance of these items. The standardisation is all the more remarkable when we consider that these weapons would have been made manually, by hot forging bloomery iron of different composition and hardness on an anvil. Bearing in mind that each of these represents an individual production event, cumulative copy-errors could have led to much larger degrees of variability, even in a short period of time. The fact that these are not documented suggests that only one, or very few, lances, were used as models for subsequent generations. Only a subgroup of six Nydam lances, accounting for the extreme outliers, displayed significant variation from the mean dimensions and shape. These might therefore constitute imitations from a different workshop. As this study was based on lances from the main deposits in Jutland, it will be interesting in the future to compare these results to those from lances found in graves outside this area.



**Fig. 14.** PCA biplot of lance shapes (GPA aligned co-ordinates) showing warpgrids of most extreme cases along PC1 (indicated by arrows, where the left warpgrid relates to the most negative outlier along PC1, and right warpgrid to the most positive outlier along PC1).

**Table 2**

ANOVA for and analysis of Havor lance shape symmetry ( $n = 77$ ) on the basis of 7 landmarks (three pairs, one on the median plane). Individual effect ('ind') represents variation between individual lances (symmetric component), 'side' represents DA (differences between original and mirrored copy for each lance).

	DF	SS	MS	Rsq	F	Z	Pr(> F)
ind	77	1.55153	0.0201498	0.99820	577.3822	23.2156	0.001
side	1	0.00011	0.0001059	0.00007	3.0349	1.3652	0.063
Residuals	77	0.00269	0.0000349	0.00173	–	–	–

## 8.2. Limitations of the metric approach

The results from metric analysis have been effective in depicting a highly standardised weapon. The study of CV, although insightful, limits the focus to any single feature at one time. CVs can be compared, though this remains to be a comparison of individual parts. In any case, as argued above, the metric approach to measuring standardisation may not be appropriate for analysing artefacts from a pre-industrial context, where standardisation might have been perceived differently.

The metric approach has been developed here to include two-way variable interactions, studying Pearson  $r$  as a means to avoiding studying any single variable in isolation. Although correlations can be defined, and thus related elements of a standardised schema identified, the approach remains limited to discussing a pair of variables at any one time. Even if all dimensions recorded are studied collectively using multivariate methods, a conceptual flaw still remains. That main flaw is that the dimensions recorded are essentially arbitrary and reflective of

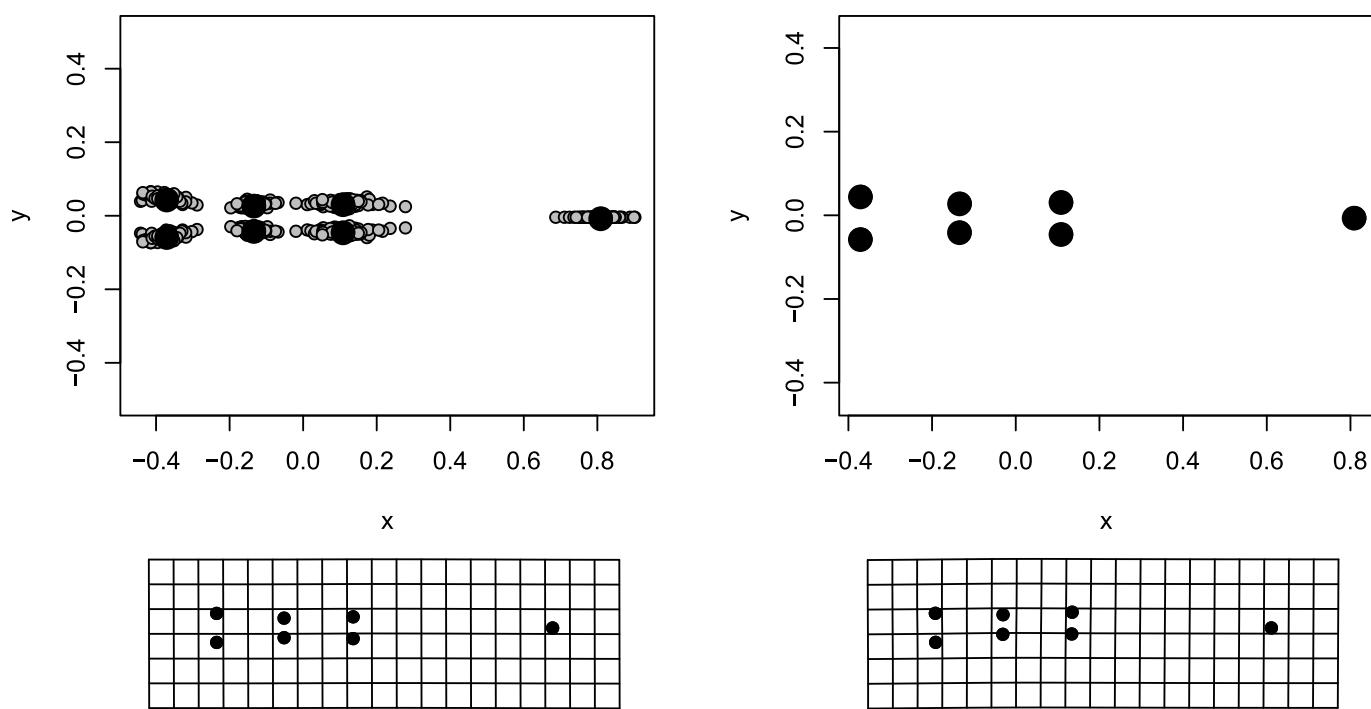
researcher choice, and may not be congruent with those relevant to the original producer or intended consumer.

## 8.3. Comparison of metric and GMM data for multivariate statistical analysis

Both datasets highlight the importance of the width dimensions in relation to the lengthwise axis of the Havor lance, and both datasets confirm that little distinction can be made between Havor lances based on site affiliation. The total variation explained in PCA by GMM data (98%) is a significant improvement on that by using metric data (70%), confirming that shape is a valid means of investigating and describing weapon standardisation. The implementation of warpgrids also provides a visual means of describing any variation observed in relation to the mean shape. As has been noted, the PCA biplot arising from the GMM data discriminates two groups within the Nydam outliers, which were otherwise clustered together in the PCA biplot using metric data.

## 8.4. Advantages of GMM analysis

The advantage of shape analysis is that the total appearance as a whole is considered, rather than artificially promoting any particular feature(s) or dimension(s) as a defining characteristic of standardisation. Testing mean shape difference can be undertaken by studying multivariate means, variation within groups, or variation between groups. Thus, two-way statistical testing of mean shape difference by multiple methods (Hotelling  $T^2$  statistic, Goodall F test and Procrustes ANOVA) ensures that all results and derived conclusions are robust. GMM analysis also allows for asymmetry to be quantified and analysed



**Fig. 15.** Symmetric component (top left) and asymmetric component (top right) of Havor lance shape, accompanied by warpgrids of mean directional asymmetry (bottom left) and mean fluctuating asymmetry (bottom right).

**Table 3**

Summary table of the relative difference (%) between x and y values for each landmark pair, expressed as the median and mean differences, as well as showing the standard deviation (StDev) of those differences.

(n = 78)	Socket Landmarks 1 and 7		Neck Landmarks 2 and 6		Width Landmarks 3 and 5	
	x	y	x	y	x	y
Median (%)	0.18	3.98	0.53	3.75	0.73	5.44
Mean (%)	0.21	4.81	0.64	5.41	1.48	7.47
StDev	0.18	4.63	0.54	5.36	2.86	6.22

for variance. It is realised, however, that only three groups (archaeological sites) were specified in this study, and thus extrapolating such tests to larger group sizes may require further modifications (or different use) of statistical methods. The multi-faceted statistical approach was only made possible through the collection and analysis of GMM data, regarded as a methodological improvement on the collection and analysis of metric data.

It should be recognised that the metric dimensions and GMM landmarks recorded in this study reflect individual preferences and opportunistic approaches to the data available. In particular, future work may also consider the adoption of 3D recording, as well as the use of landmark-free comparisons of shape, whether 2D or 3D (cf. Bevan et al., 2014). The allometric relationship identified in this analysis would also be worthy of further investigation. These will no doubt provide more comprehensive assessments of variation in size and shape, though an explicit cost-benefit analysis may be appropriate.

#### 8.5. Weapon symmetry

Using the GMM data, it was possible to demonstrate that there was no significant difference overall in Havor lance shape symmetry, confirming both reflected sides to be the same. Further inspection of the asymmetrical GMM data made it possible to closely examine differences between landmark co-ordinate pairs, which should be zero for a

perfectly symmetrical shape. The mean and median differences border on the line of human copy error and thus confirm that symmetry was deliberately intended for the Havor lance shape. It is expected that the approaches employed here to quantify and compare artefact symmetry may be of use in other fields – for example in lithic analysis, where symmetry often constitutes a subject of interest.

#### 9. Conclusion: advocating 'shape' analysis as a standard tool for monitoring weapon standardisation

The case study presented here employed shape analysis and compared it to traditional metric approaches. The limitations of the traditional metric approach have been highlighted and the advantages of GMM analysis described. Shape analysis has been demonstrated to be an effective tool for investigating variation, used indirectly as a measure for object standardisation. GMM analysis has improved greatly on metric methods by yielding robust and improved results for identifying and statistically assessing standardisation of the total appearance and not any one part. The analysis of mean shape difference can thus identify and empirically demonstrate products to be highly standardised in terms of their shape variation, as has been demonstrated here for the Havor lance, as well as allowing comparisons between and within sites. The results from this study advocate that shape analysis should become a standard tool of investigation that complements

traditional metric analysis and studies of variance as a means to investigating and understanding weapon standardisation using statistical methods.

## Author contributions

Birch conducted the analyses and wrote the main draft of the manuscript, which was jointly revised with input from Martinón-Torres on standardisation, symmetry and archaeological interpretation.

## Acknowledgements

This work is based in part on PhD research in the Department of Archaeology, University of Aberdeen. Special thanks are made to Peter Crew, Mike Charlton, Keith Dobney and Thilo Rehren for reading the chapter form of this work, and to Allowen Evin and Christian Hoggard for their comments on a draft of this paper. Birch wishes to dedicate this paper to Garðar Guðmundsson, with whom many long and fruitful discussions were had over the Havor lance. The final write-up of this research was supported by funding from The Danish National Research Foundation is the funding source, and the grant is DNRF119, which is for the Centre of Excellence for Urban Network Evolutions (UrbNet), where I work. The authors accept responsibility for any errors or omissions made.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jas.2018.11.002>.

## References

- Adams, D.C., Otárola-Castillo, E., 2013. geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* 4, 393–399.
- Archer, W., Braun, D.R., 2010. Variability in bifacial technology at Elandsfontein, Western cape, South Africa: a geometric morphometric approach. *J. Archaeol. Sci.* 37, 201–209. <https://doi.org/10.1016/j.jas.2009.09.033>.
- Bemmern, G., Bemmern, J., 1998a. Der Opferplatz von Nydam: Die Funde aus den älteren Grabungen: Nydam-I und Nydam-II. Band 1: Text. Wachholz Verlag, Neumünster.
- Bemmern, G., Bemmern, J., 1998b. Der Opferplatz von Nydam: Die Funde aus den älteren Grabungen: Nydam-I und Nydam-II. Band 2: Katalog und Tafeln. Wachholz Verlag, Neumünster.
- Bevan, A., Li, X., Martinón-Torres, M., Green, S., Xia, Y., Zhao, K., Zhao, Z., Ma, S., Cao, W., Rehren, T., 2014. Computer vision, archaeological classification and China's terracotta warriors. *J. Archaeol. Sci.* 49, 249–254. <https://doi.org/10.1016/j.jas.2014.05.014>.
- Bignon, O., Baylac, M., Vigne, J.-D., Eisenmann, V., 2005. Geometric morphometrics and the population diversity of Late Glacial horses in Western Europe (*Equus caballus arcelini*): phyleogeographic and anthropological implications. *J. Archaeol. Sci.* 32, 375–391. <https://doi.org/10.1016/j.jas.2004.02.016>.
- Birch, T., 2018. Standardised manufacture of iron Age weaponry from southern Scandinavia: constructing and provenancing the havor lance. In: Dolfini, A., Crellin, R.J., Horn, C., Uckelmann, M. (Eds.), Prehistoric Warfare and Violence: Quantitative and Qualitative Approaches, Quantitative Methods in the Humanities and Social Sciences. Springer International Publishing, Cham, pp. 247–276. [https://doi.org/10.1007/978-3-319-78828-9\\_12](https://doi.org/10.1007/978-3-319-78828-9_12).
- Bronstet, J., 1960. Danmarks Oldtid III. Jernalderen, second ed. Glyndal, København.
- Cucchi, T., Hulme-Beaman, A., Yuan, J., Dobney, K., 2011. Early Neolithic pig domestication at Jiahu, Henan Province, China: clues from molar shape analyses using geometric morphometric approaches. *J. Archaeol. Sci.* 38, 11–22. <https://doi.org/10.1016/j.jas.2010.07.024>.
- Curran, S.C., 2012. Expanding ecomorphological methods: geometric morphometric analysis of *Cervidae* post-crania. *J. Archaeol. Sci.* 39, 1172–1182. <https://doi.org/10.1016/j.jas.2011.12.028>.
- de Azevedo, S., Charlin, J., González-José, R., 2014. Identifying design and reduction effects on lithic projectile point shapes. *J. Archaeol. Sci.* 41, 297–307. <https://doi.org/10.1016/j.jas.2013.08.013>.
- Dryden, I.L., 2013. Shapes: Statistical Shape Analysis. R Package Version 1.1-8. [WWW Document]. <http://CRAN.R-project.org/package=shapes>.
- Eerkens, J.W., 2000. Practice makes within 5 % of perfect: the role of visual perception, motor skills, and human memory in artifact variation and standardization. *Curr. Anthropol.* 41, 663–668.
- Eerkens, J.W., Bettinger, R., 2001. Techniques for assessing standardization in artifact assemblages: can we scale material variability? *Am. Antiqu.* 66, 493–504.
- Eerkens, J.W., Lipo, C.P., 2005. Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological records. *J. Anthropol. Archaeol.* 24, 316–334.
- Engelhardt, C., 1865. Nydam Mosefund 1859–1863. Gad, København.
- Evin, A., Cucchi, T., Cardini, A., Strand Vidarsdóttir, U., Larson, G., Dobney, K., 2013. The long and winding road: identifying pig domestication through molar size and shape. *J. Archaeol. Sci.* 40, 735–743. <https://doi.org/10.1016/j.jas.2012.08.005>.
- Goodall, C.R., 1991. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. Ser. B* 53, 285–339.
- Gower, J.C., 1975. Generalized procrustes analysis. *Psychometrika* 40, 33–51.
- Hansen, U.L., 2007. Barbarians in the north - the greatest concentrations of roman weaponry in Europe. In: Grane, T. (Ed.), Beyond the Roman Frontier: Roman Influences on the Northern Barbaricum. Edizioni Quasar di Severino Tognon srl, Rome, pp. 105–130.
- Hansen, U.L., 2003. 150 years of weapon-offering finds - research and interpretations. In: Jorgensen, L., Storaagaard, B., Thomsen, L.G. (Eds.), The Spoils of Victory: the North in the Shadow of the Roman Empire. Nationalmuseet, København, pp. 84–89.
- Humphris, J., Martinón-Torres, M., Rehren, T., Reid, A., 2009. Variability in single smelting episodes - a pilot study using iron slag from Uganda. *J. Archaeol. Sci.* 36, 359–369.
- Ilkjær, J., 2000. Illerup Ådal - Archaeology as a Magic Mirror. Jysk Arkæologisk Selskab, Moesgård.
- Ilkjær, J., 1990a. Illerup Ådal: Die Lanzen und Speere. Textband. Jysk Arkæologisk Selskab, Moesgård.
- Ilkjær, J., 1990b. Illerup Ådal: Die Lanzen und Speere. Tafelband. Jysk Arkæologisk Selskab, Moesgård.
- Ilkjær, J., Lønstrup, J., 1982. Interpretation of the great votive deposits og iron Age weapons. *J. Dan. Archaeol.* 1, 95–103.
- Imer, L., 2010. Runes and romans in the north. *Futhark Int. J. Runic Stud.* 1, 41–64.
- Klingenberg, C.P., 2015. Analyzing fluctuating asymmetry with geometric morphometrics: concepts, methods, and applications. *Symmetry* 7, 843–934. <https://doi.org/10.3390/sym7020843>.
- Li, X.J., Bevan, A.H., Martinón-Torres, M., Rehren, T., Cao, W., Xia, Y., Zhao, K., 2014. Crossbows and imperial craft organisation: the bronze triggers of China's Terracotta Army. *Antiquity* 88, 126–140.
- Lycett, S.J., Cramon-Taubadel, N. von, Gowlett, J.A.J., 2010. A comparative 3D geometric morphometric analysis of Victoria West cores: implications for the origins of Levallois technology. *J. Archaeol. Sci.* 37, 1110–1117. <https://doi.org/10.1016/j.jas.2009.12.011>.
- Lycett, S.J., von Cramon-Taubadel, N., 2013. A 3D morphometric analysis of surface geometry in Levallois cores: patterns of stability and variability across regions and their implications. *J. Archaeol. Sci.* 40, 1508–1517. <https://doi.org/10.1016/j.jas.2012.11.005>.
- Martinón-Torres, M., Bastir, M., Bermúdez de Castro, J.M., Gómez, A., Sarmiento, S., Muñoz, A., Arsuaga, J.L., 2006. Hominin lower second premolar morphology: evolutionary inferences through geometric morphometric analysis. *J. Hum. Evol.* 50, 523–533. <https://doi.org/10.1016/j.jhevol.2005.12.004>.
- Martinón-Torres, M., Li, X.J., Bevan, A., Xia, Y., Zhao, K., Rehren, T., 2014. Forty thousand arms for a single emperor: from chemical data to the labor organization behind the bronze arrows of the terracotta army. *J. Archaeol. Method Theor.* 21, 534–562. <https://doi.org/10.1007/s10816-012-9158-z>.
- Odler, M., 2016. Old Kingdom Copper Tools and Model Tools, Archaeopress Egyptology. Archaeopress, Oxford.
- Okumura, M., Araujo, A.G.M., 2014. Long-term cultural stability in hunter-gatherers: a case study using traditional and geometric morphometric analysis of lithic stemmed bifacial points from Southern Brazil. *J. Archaeol. Sci.* 45, 59–71. <https://doi.org/10.1016/j.jas.2014.02.009>.
- Ørsnes, M., 1988. Ejbsøl I: Waffenopferfunde des 4.–5. Jahrh. nach Chr. Det Kongelige Nordiske Oldskriftselskab, København.
- Owen, J., Dobney, K., Evin, A., Cucchi, T., Larson, G., Strand Vidarsdóttir, U., 2014. The zooarchaeological application of quantifying cranial shape differences in wild boar and domestic pigs (*Sus scrofa*) using 3D geometric morphometrics. *J. Archaeol. Sci.* 43, 159–167. <https://doi.org/10.1016/j.jas.2013.12.010>.
- Pauli Jensen, X., 2010. Som brødre vi dele - ny tolkning af Vimoses våbenofre. *Natl. Arb.* 2010, 23–35. [https://www.academia.edu/1479168/Som\\_br%C3%88dre\\_vi\\_dele\\_-ny\\_tolkning\\_af\\_Vimoses\\_v%C3%A5benofre](https://www.academia.edu/1479168/Som_br%C3%88dre_vi_dele_-ny_tolkning_af_Vimoses_v%C3%A5benofre).
- Pauli Jensen, X., 2009. From fertility rituals to weapon sacrifices: the case of the south Scandinavian bog finds. In: von Freeden, U., Friesinger, H., Wamers, E. (Eds.), Glaube, Kult Und Herrschaft: Phänomene Des Religiösen Im 1. Jahrtausend N. Chr. In Mittel- Und Nordeuropa. Dr. Rudolf Habelt GmbH, Bonn, pp. 53–64.
- Petersen, P., 1995. Nydam Offermose: Mosefund i Damnak gennem 2000 år - og de store offerfund i moserne som grundlag for en fyldigere beskrivelse af Nydam offermose. Dansk Historisk Håndbogsforlag, København.
- QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation.
- R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Ros, J., Evin, A., Bouby, L., Ruas, M.-P., 2014. Geometric morphometric analysis of grain shape and the identification of two-rowed barley (*Hordeum vulgare* subsp. *distichum* L.) in southern France. *J. Archaeol. Sci.* 41, 568–575. <https://doi.org/10.1016/j.jas.2013.09.015>.
- Roux, V., 2003. Ceramic standardization and intensity of production: quantifying degrees of specialization. *Am. Antiqu.* 68, 768–782. <https://doi.org/10.2307/3557072>.
- Seetha, K., Cucchi, T., Dobney, K., Barker, G., 2014. A geometric morphometric re-evaluation of the use of dental form to explore differences in horse (*Equus caballus*) populations and its potential zooarchaeological application. *J. Archaeol. Sci.* 41, 904–910. <https://doi.org/10.1016/j.jas.2013.10.022>.

- Stark, B., 1995. Problems in analysis of standardization and specialization in pottery. In: Mills, B.J., Crown, P.L. (Eds.), *Ceramic Production in the American Southwest*. University of Arizona Press, Tucson, pp. 231–267.
- Stegmann, M.B., Gomez, D.D., 2002. *A Brief Introduction to Statistical Shape Analysis (Technical Report). Informatics and Mathematical Modelling*. Technical University of Denmark, DTU, Denmark.
- Terral, J.-F., Alonso, N., Capdevila, R.B. i, Chatti, N., Fabre, L., Fiorentino, G., Marinval, P., Jordá, G.P., Pradat, B., Rovira, N., Alibert, P., 2004. Historical biogeography of olive domestication (*Olea europaea* L.) as revealed by geometrical morphometry applied to biological and archaeological material. *J. Biogeogr.* 31, 63–77. <https://doi.org/10.1046/j.0305-0270.2003.01019.x>.
- Wilczek, J., Monna, F., Barral, P., Burlet, L., Chateau, C., Navarro, N., 2014. Morphometrics of Second Iron Age ceramics – strengths, weaknesses, and comparison with traditional typology. *J. Archaeol. Sci.* 50, 39–50. <https://doi.org/10.1016/j.jas.2014.05.033>.
- Wilczek, J., Monna, F., Gabillot, M., Navarro, N., Rusch, L., Chateau, C., 2015. Unsupervised model-based clustering for typological classification of Middle Bronze Age flanged axes. *J. Archaeol. Sci. Rep.* 3, 381–391. <https://doi.org/10.1016/j.jasrep.2015.06.030>.