International Journal of Osteoarchaeology

WILEY

RESEARCH ARTICLE OPEN ACCESS

Assessing the Greater Sciatic Notch With 2-D Shape Analysis for Sex Estimation

Samuel R. Rennie[1] | Heather M. Tamminen[2]

[1]Department of Life and Environmental Sciences, Faculty of Science and Technology, Bournemouth University, Poole, UK | [2]Department of Archaeology and Anthropology, Faculty of Science and Technology, Bournemouth University, Poole, UK

Correspondence: Samuel R. Rennie (srennie@bournemouth.ac.uk)

Received: 7 August 2024 | **Revised:** 8 January 2025 | **Accepted:** 10 January 2025

Keywords: biological profile | computational methods | elliptic Fourier analysis | greater sciatic notch | morphometric methods | *os coxae* | sex estimation | shape analysis

ABSTRACT

A correct biological profile leads to a better understanding of the past and assists in the identification of human remains within bioarchaeology and forensic casework. Sex estimation forms a critical component of a biological profile. With the advancement of technologies such as geometric morphometrics (GMM), new methods and a deeper understanding of morphological features can be investigated digitally. However, how well do these methods compare to standard visual methods and how easy are they to employ? This research investigates the use of 2-D shape analysis and visual morphological methods for sex estimation using the greater sciatic notch (GSN). A total of 202 adult *os coxae* were photographed and analyzed from the Spitalfields Coffin Plate Collection housed at the Natural History Museum, UK. Each *os coxae* was analyzed digitally to extract a "line" for elliptical fourier analysis (EFA) and subsequent discriminant function analysis (DFA). *Os coxae* were also scored using two well established morphological methods for the GSN. This study found an overall accuracy of 82.81% when using the computational method (EFA and DFA). Lower accuracies were found for the visual methods with the Bruzek method correctly classifying 82.17% and the Walker method resulting in a much lower accuracy at 72.77%. The finding of this study showcases the benefits of using more computational methods such as shape analysis/GMM. However, it has a nearly identical overall error rate to the Bruzek method and higher accuracy than the Walker method and therefore is a suitable and accurate method for sex estimation. As these practices are evolving, practitioners will have to balance the cost/benefit (e.g., time, training, and accuracy) of using the different techniques while continuing to refine and combine approaches for optimal results in biological profiling.

1 | Introduction

Estimating a biological profile is the culmination of different analyses including sex, age-at-death, stature, and ancestry estimations. All of these aspects are used when investigating both forensic and archaeological human remains, in situations such as human identification and population-level demographics.

Sex estimation relies on understanding the morphological variation between the sexes and being able to differentiate between sexually dimorphic shapes (Phenice 1969; Klales 2020). Correctly identifying the sex of an individual helps build a more informative profile of human populations and human demographic structures of the past as well as assisting in the identification of unknown remains within forensic cases.

Sexing methods across the human skeleton include metric analyses (Asala 2001, 2002; Murphy 2005; Harma and Karakas 2007; Kranioti et al. 2009; Brůžek et al. 2017; Cuzzulin et al. 2020; Maio, Cunha, and Navega 2024), morphological analyses (Acsádi and Nemeskéri 1970; Bruzek 2002; Walker 2005; Klales, Ousley, and Vollner 2012; Rennie, Eliopoulos, and Gonzalez 2023), and computational techniques such as geometric morphometric analyses (GMM) (Franklin et al. 2012; Bertsatos et al. 2020; Gillet et al. 2020; Conner et al. 2024). Metric analysis relies on overall size dimorphism, which, in humans, is characterized as males being generally larger than females in overall body size and most skeletal elements. However, this pattern is not necessarily followed in the pelvic region, where females exhibit some larger dimensions, particularly in the areas associated with childbirth (e.g., pubic length and pelvic inlet) (Correia, Balseiro, and De Areia 2005; Kurki 2007, 2011, 2013a, 2013b). The localized dimorphism of the pelvic region develops early in fetal life and continues into adulthood and is primarily attributed to evolutionary pressures of childbirth (Gunstra et al. 2023; Kanahashi et al. 2024).

GMM and standard morphological analyses rely on differences in overall shape, size, and form of skeletal features. The skull and the *os coxae* are the two major elements biological anthropologists focus on when observing sexual dimorphism visually. Walker (2008) is a very commonly used method for sex estimation using the skull, which is based on observing five features and scoring them across a five-point grading system. The most famous and well used method for sex estimation using the *os coxae* is the Phenice method, which was published in 1969, and which looked at three aspects of the o*s coxae*. It was modernized in 2012 by Klales, Ousley, and Vollner (2012) who changed the original scoring of "male - indeterminate - female" categories to a "1–5" grading system similar to Walker in 2008.

The sexually dimorphic greater sciatic notch (GSN) of the *os coxae* was described to be typically narrow in males and wide in females (Lazorthes and Lhez 1940; Jovanonic and Zivanovic 1965; Singh and Potturi 1978). From the general description of the GSN angle, Acsádi and Nemeskéri (1970) created a 5-point grading system to assess the sexual dimorphism between males and females and integrated it into their "complex method" for sex estimation. From there, the European Meeting of Anthropologists (1980) and Buikstra and Ubelaker (1994) adapted Acsádi and Nemeskéri's (1970) descriptions in an attempt to standardize osteological techniques for sex estimation. Walker (2005) published his results on the descriptions from Buikstra and Ubelaker (1994) and found that 88% of females exhibited the extreme morphology of "Grade 1," while males showed more variation with 91% ranging from the intermediate score of "3" to the hyper masculine score of "5." Gómez-Valdés et al. (2012) applied the Walker (2005) descriptions on a modern Mexican population to find equivalent results. Most females were deemed to be in the hyper feminine score of 1 while males showed more variation and spanned the ordinal scale.

From applying visual methods, researchers have attempted to "metricise" the GSN by calculating the angle and linear measurements (Singh and Potturi 1978; MacLaughlin and Bruce 1990). Researchers have also investigated its morphology using GMM; however, they stated that further work needs to be done for this method to become normal practice (Gómez-Valdés et al. 2012; Velemínska et al. 2013). One of the issues with this approach is the length of time it takes to analyze each specimen; hence the "classic" visual technique of scoring is preferred in both archaeological and forensic cases. The morphology of the GSN is also one of the preferred sexing techniques because of it robusticity and likelihood of survival. Waldron (1987) and Stojanowski, Siedmann, and Doran (2002) observed the rate of preservation of skeletal elements from different archaeological assemblages and found that when focusing on the pelvic bones, the sacroiliac joint, including the adjacent GSN, was preserved at a much higher frequency than the *os pubis*.

Recently, several studies have examined the use of computational methods to estimate the sex of human remains, often with the goal of higher accuracy, less subjectivity, and faster implementation. Studies such as Kilmer and Garvin (2020) have explored shape analysis in the GSN and Conner et al. (2024) have examined the use of landmarks and GMM analysis. Both demonstrate promising results and call for further analysis on the topic.

As sex estimation techniques continue to evolve, practitioners will have to balance the costs and benefits (e.g., time required, length of training, and accuracy of methods) of using different methods while continuing to refine and combine approaches for the most accurate results in biological profiling. The aim of this article is to better understand the application of 2-D shape analysis in sex estimation using the GSN and to build on this body of knowledge. It investigates whether using a more computational approach using open-source code yields more accurate results than the standard morphological methods for assessing sex.

2 | Method and Materials

Human remains from the Spitalfields Coffin-Plate collection housed at the Natural History Museum in London (UK) were used for this analysis. The collection is dated between the 18th and 19th centuries and was originally excavated from the church grounds and vaults in 1984 (Molleson and Cox 1993; Reeves and Adams 1993). From the excavation, 968 individuals were found in total with 387 having coffin plates associated with them. The coffin plates contained information on those interred such as name, age, and year of death. This information coupled with archival research of the parish records helped confirm the age and sex of the individuals (Molleson and Cox 1993, Reeves and Adams 1993).

For individuals to be selected, only those with an undamaged sacroiliac joint aged 18 years and above were chosen (Rennie 2018). Of the 387 individuals from the known age and sex collection, 202 (106 female, 96 male) *os coxae* were photographed. The selected *os coxae* had a mean age of 57 years (min = 21 years old, max = 91 years old). The remaining 185 Individuals were excluded due to taphonomic damage to their sacroiliac area or pathology that affected the shape of their *os coxa*e. Ethical approval was obtained prior to data collection.

Each *os coxae* was photographed using a Canon EOS Rebel T3 1100D. The photography set up followed a standard procedure outlined in Velemínska et al. (2013). The camera was set on a tripod at a 40 cm height with the lens facing perpendicular (90°) to the table. The *os coxa* was laid flat on the iliac blade with the *os pubis* facing toward the lens with a "L" scale bar. From each photograph, the shape of GSN was digitized using Adobe Illustrator to create a line running from the retroarticular area (piriform tubercle) to the tip of ischial spine, along the GSN (following methods by Kilmer and Garvin 2020, p3). The Left GSN was given preference when it was present; however, in situations where no left GSN was available, the right was used and mirrored (Walker 2005). A single observer captured the images (SRR) and digitized the GSN outlines (HMT).

Initially, the GSNs were outlined using a similar procedure to Kilmer and Garvin (2020), as a solid black shape. However, during the preliminary analysis, it was found that the "imaginary" line connecting the top and bottom of the GSN was influencing the analysis by warping the shape when creating the outline to run the EFA. This created an ovoid shape which did not visually appear to represent the GSNs. As a result, only the outline of the GSN was used (Figure 1). This issue does not seem to have been encountered by Kilmer and Garvin (2020) with the software that they used and draws attention to the fact that even computational methods are not all the same and care must be taken to avoid introducing error.

Once digitized, statistical analyses were undertaken in the R environment to analyze the shapes extracted from the GSNs (Hoggard 2020). For this, an Elliptical Fourier Analysis (EFA) using the R package "Momocs" (Bonhomme et al. 2014) was performed. EFA traditionally outlines shapes/areas of interest rather than lines; therefore, the authors created an area of interest by outlining firstly from the piriform tubercle down to the ischial spine and then in reverse without overlapping to create a contour that is at least two pixels thick (Figure 1). Once the digital line was created, a Generalized Procrustes Analysis (GPA) was performed which removes factors of size, rotation, and orientation (Klingenberg 2011). The EFA then produced a visual depiction of the common shape changes, showing the variation within the sample. Age-at-death has previously been found to be a non-significant covariate for this type of analysis and therefore was excluded here (Kilmer and Garvin 2020). Secular change has been noted in the morphology of pelvic traits described in Klales, Ousley, and Vollner (2012), and even though it was statistically found, Klales (2016) states that the three aspects of the anterior *os coxae* are still applicable from historic to modern day human remains.

In addition to performing the 2D shape analysis, each of the *os coxa* were visually scored and sex was assessed using the Walker (2005) method, and the greater sciatic notch aspect of

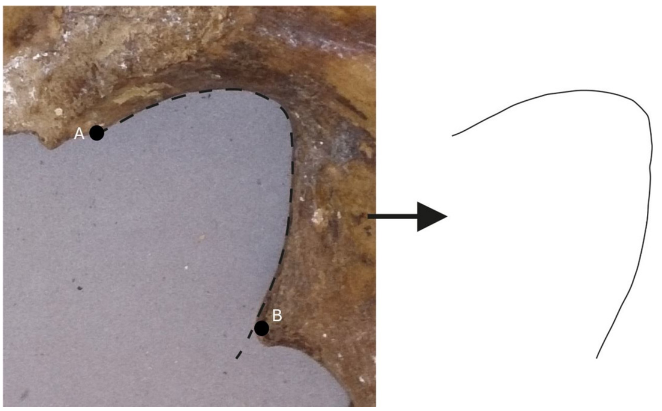

FIGURE 1 | An example of a greater sciatic notch and the resultant line. Point A is the piriform tubercle and Point B is the ischial spine. [Colour figure can be viewed at wileyonlinelibrary.com]

Bruzek's (2002) complex method. The *os coxae* were scored before the computational analysis was performed. For the Walker Method (2005), the *os coxa* was aligned so the GSN matched the orientation of the GSNs in the reference image and following the description noted on p.386 of Walker (2005). In the present study, the authors designated a score of "2" as female morphology and a score of "3" as indeterminate morphology. These adaptations to Walker (2005) helped facilitate a more straightforward comparison with other published studies that have adopted similar adjustments in their classifications (Table 1) (Gómez-Valdés et al. 2012; Stinnesbeck et al. 2020; Conner et al. 2024; Russell et al. 2024). In this study, Bruzek's (2002) GSN character was tested, where each of the three conditions were observed and recorded (Condition 1: Proportion of length of the sciatic notch chords; Condition 2: form of the contour notch chords; and Condition 3: contour of the posterior notch chord relative to line from point "A" to sciatic notch breadth, Bruzek 2002, 159–161).

To enhance comparability with existing literature, the Bruzek method (2002) was adapted to align with Walker's approach (2005) by assigning numerical values to each combination of conditions (Table 1) (Walker 2005, 2008; Klales, Ousley, and Vollner 2012; Rennie, Eliopoulos, and Gonzalez 2023). All conditions from the Bruzek method (2002) for the GSN were treated as having equal weighting for estimating sex.

Intraobserver and interobserver error rates were calculated for the visual morphological methods by computing a linear weighted kappa Cohen's value and then using Landis and Koch (1977) descriptions for interpretation. For this aspect, 40 randomly selected *os coxae* were used. Intraobserver analysis was performed 2 weeks after initial data collection. An observer with relevant expertise was selected to conduct the interobserver analysis. They were a PhD candidate in biological anthropology at the time of data collection and had 4 years of experience analyzing human skeletal remains, with a focus on cranial non-metric traits.

3 | Results

After performing the EFA, two main harmonics were apparent. The generated images with negative values are shown to be male, and those with positive to be female. From this visualization, female morphology is mainly focused on the overall depth of the notch between the piriform tubercle and the ischial spine (Figures 2 and 3, PC1).

The differences in shape between the GSN were found to be primarily accounted for by the first two principal components, with minor contributions of the third. The first Principal Component

TABLE 1 | Conversion of Bruzek (2002) responses for the GSN into a 1–5 grading system, similar to Walker (2005).

Female	Probable female	Indeterminate	Probable male	Male
FFF	FFM/FFI	MFI/MII/FII/III	MMI/MMF	MMM
1	2	3	4	5

Abbrevitaions: F = female; I = indeterminate; M = male.

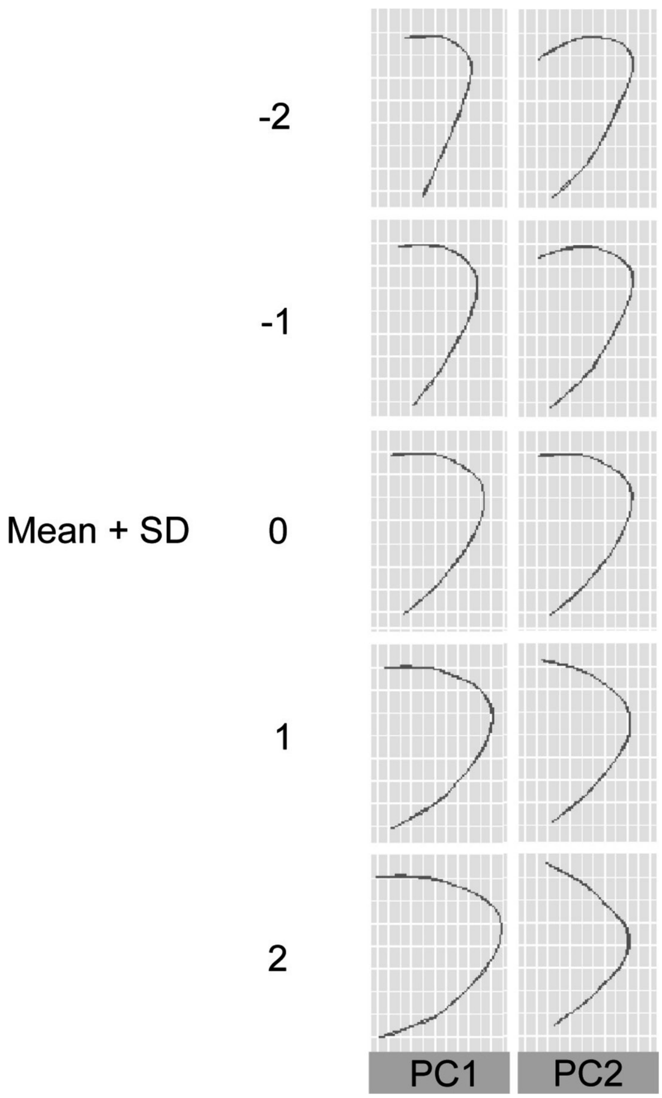

FIGURE 2 | Visualization of the shape variation of the GSN along axes PC1 and PC2.

(PC1) accounts for 53.7% of the variation within the sample with PC2 accounting for 37.4%.

Clear separation is noted between the female and male specimens in the sample (Figure 3). Although PC1 accounted for the highest percentage of variation, it was PC2 that demonstrated the most success in discriminating between male and female *os coxae*. Using a cross-validated LDA, it was found that the rate of correct classification of the male and female GSNs was approximately 82.81% (female: 84.91%, male: 79.17%).

Each of the 202 *os coxae* were visually scored using the Walker (2005) method and the GSN descriptions by Bruzek (2002). *Os coxae* that were scored as indeterminate (a score of "3") were considered an incorrect classification. The Walker (2005) method yielded the lowest overall and male accuracy, however, performed better with correctly classifying female *os coxae*. Bruzek (2002), on the other hand, had a slightly better accuracy for male classification than female, and yielded a much more balanced result (Tables 2 and 3).

Intraobserver and interobserver error rates were calculated for the Walker (2005) and Bruzek (2002) methods using 40 randomly selected *os coxa*. Results from the kappa Cohen's test show substantial agreement was found for the intraobserver error rates for both Walker (2005) ($k = 0.690$, $p < 0.001$) and Bruzek (2002) ($k = 0.721$, $p < 0.001$). When looking at interobserver error rates, there was substantial agreement for both Walker (2005) ($k = 0.650$, $p < 0.001$) and Bruzek (2002) ($k = 0.705$, $p < 0.001$).

4 | Discussion

This study sought to increase understanding of the sexual dimorphism of the greater sciatic notch using 2-D shape analysis and assess whether a more computational approach to sex estimation yields more accurate results than standard morphological methods for assessing sex. Overall, the 2-D shape analysis

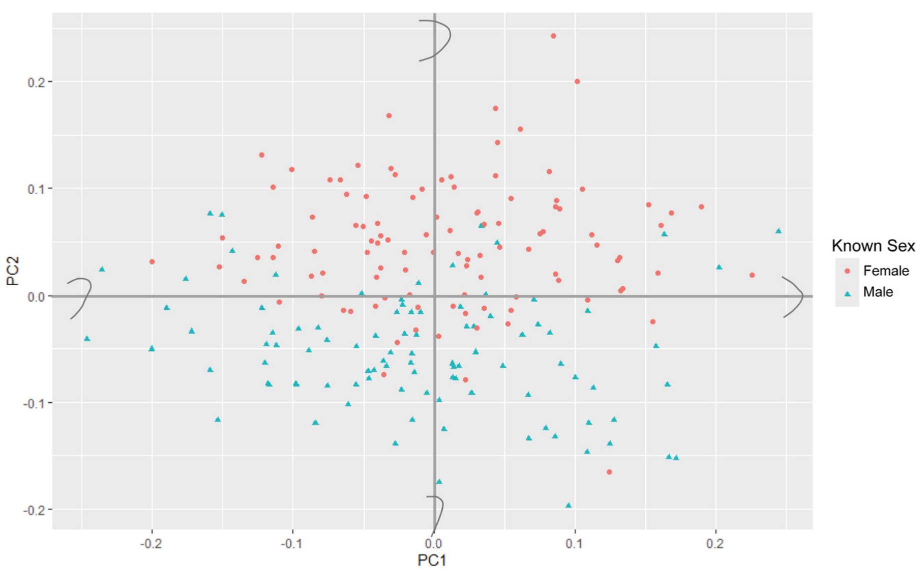

FIGURE 3 | A principal component analysis biplot showing PC1 and PC2. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 | The number of individuals given each classification for the visual methods of Walker (2005) and Bruzek (2002).

Walker (2005)					
	1	2	3	4	5
Male	7	6	23	31	29
Female	68	24	9	2	3
Bruzek (2002)					
	FFF	FFI/ FFM	FII/FMI/ III/MII	MMI/ MMF	MMM
Male	12	0	3	0	81
Female	87	0	7	0	12

TABLE 3 | The percentage of correct classification across male, female, and overall using Walker (2005), Bruzek (2002), and this study.

Method	% Male correct	% Female correct	% Overall	% Sex difference
Walker 2005	62.50	86.79	75.23	−24.29
Bruzek 2002	84.38	82.08	82.17	2.30
2D shape analysis	79.17	84.91	82.81	−5.74

yielded very similar results to the Bruzek (2002) method, and outperformed Walker (2005) (Table 3). However, the method that performed the best for this study in terms of overall accuracy and having the lowest sex bias in results was Bruzek (2002).

Kilmer and Garvin (2020) used outline analysis/EFA to gain a better understanding of the GSN and the obturator foramen in terms of sexual dimorphism. Overall, the results from this study regarding the GSN (82.81%) were not dissimilar to Kilmer and Garvin (2020) who recorded accuracies of 86.80% (pooled). However, differences between the current study and Kilmer and Garvin (2020) occurred when observing male and female accuracies separately. Kilmer and Garvin (2020) found that the outline analysis performed better for males (89.70%) than females (83.70%), while in this study, it was the inverse with males recording a lower accuracy (79.17%, Table 3). These potential differences could be explained by having to adapt the method Kilmer and Garvin (2020) had used for their study. Creating a solid black shape when connecting the iliac spine to the posterior portion of the preauricular area produced results that massively distorted the shape of the GSN with the software used in the present study; therefore, just the line drawn along the edge of the bone was used. Additionally, the PC descriptions between this study and Kilmer and Garvin (2020) appear to be generally similar, though their use of multiple populations means that there are some minor differences. This is suggestive of a level of consistency across methods when using GMM/shape analysis. Further investigation would be needed to determine how interchangeable different statistical methods are.

Conner et al. (2024) analyzed a variety of skeletal populations focusing on the GSN. For the archaeological population; however, they expanded to use several sexing methods across the skeleton. After comparing the findings, the visual depiction of the PCs look similar between their study on the GSN and the current one. When investigating error rates, in modern populations, they found an accurate classification of 58% when classifying indeterminate scores as incorrect when using Walker (2005). However, when they observed if the sex was correct when excluding indeterminate scores, they found that accuracy was 75% using Walker (2005), 83.25% using GSN angle estimation, and 90.2% using GMM. Conner et al. (2024) present a compelling case for the use of GMM rather than Walker (2005) for both population samples they used. They do not, however, compare the visual sex estimation methods of Walker to those of Bruzek, which, as demonstrated in the present study, appear superior.

Velemínska et al. (2013) analyzed six different methods of GMM on the GSN. In comparison, this study underperformed as their EFA (Method E in their study) produced overall results of 92.14% for their pooled samples (US and Mexican samples). They found that 90.60% of males and 93.75% of females were correctly classified in their pooled samples, with a similar distribution when considering each population separately. These findings also align with the results from this study as females recorded a greater accuracy in classification.

Our results similarly correspond closely to Gómez-Valdés et al. (2012) who applied visual, angular, and GMM methods on a sample from UNAM (Mexico) and recorded large differences in sex estimation between the Walker (2005)/visual method and the GMM when applied to the GSN. The GMM resulted in an accuracy of 82.3% (which is in line with the results from this study), and 68.50% when using the Walker (2005) methods (Gómez-Valdés et al. 2012). However, the sex bias was similar between Walker (2005) and the GMM, with female classification obtaining a higher overall accuracy.

The differences recorded across studies, both in total accuracy and in sex differences, may reflect the different geographic and temporal skeletal populations used (modern and archaeological US, modern Portuguese, modern Mexican and archaeological Nubian versus post-medieval British) (Gómez-Valdés et al. 2012; Velemínska et al. 2013; Kilmer and Garvin 2020; Conner et al. 2024). This raises the often-discussed questions about the validity of using sex estimation methods on samples that are different to the populations the methods were created on (Spradley et al. 2008; Velemínska et al. 2013; Kilmer and Garvin 2020). This factor supports the concept of using bespoke 2-D methods for investigations such as sex estimation because that allows for the statistical models used to be trained on the same population that will be investigated.

When focusing on the visual methods, Rennie (2018) analyzed eight visual methods on the *os coxae* and found that using the Walker (2005) method produced mixed results across three different South African populations and found that it ranged between 75.60% and 82.38%. Bruzek (2002) found an accuracy of 67%–70% for females and 67%–80% for males when using the more objective method for GSN on a French and Portuguese sample. When applying the method for this sample, accuracy was slightly higher at 82.17% with a better accuracy for males than females (Table 3).

The use of shape analysis has the potential to yield a higher accuracy for sex classification compared to purely visual methods. However, it is difficult to employ these digital methods, such as GMM landmark analysis and shape/outline analysis, as they require considerably more training than the visual methods. Based on this and the finding of this study that Bruzek's (2002) GSN descriptions provides a near identical result, Bruzek's method may be most suitable where a more rapid identification is needed. Although, to better understand this, these comparisons should be applied to larger and more diverse populations (with known sex, age, and ancestry). Modern post-mortem CT scans could also be a resource to expand the sample size and diversity while further exploring the utility and limitations of the 2-D methods within this study. Conner et al. (2024) explored the use of CT data in the estimation of sex. They use both visual and computational methods, finding low error rates for classification using their GMM and GSN angle method. Further investigation into the use of deep machine learning, AI, and computational neural network analysis for pure visual analysis using medical imaging (CT, X-ray, and MR, for example) is an avenue that should be pursued.

The choice of method for a researcher or practitioner largely depends on the specific research objectives and the amount of time available. If the goal is to conduct a sex assessment, then the visual method is often sufficient. Notably, Lovell (1989) showed that a person's experience level and training in visual methods do not significantly impact accuracy. However, when opting for geometric morphometric analysis, more time must be allocated not only for data collection but also for training in these specialized techniques (Herzlinger, Goren-Inbar, and Grosman 2017; Liutu and Dixon 2020). Therefore, visual methods can provide rapid assessments while geometric morphometrics require a greater investment of time for more detailed and comprehensive results.

However, the computational methods allow for a population-specific analyses to be performed, as long as sufficient training data is available from that particular population. Therefore, there are instances where the extra investment in the computational methods would be more beneficial, especially when dealing with a large number of samples from a unique or less studied population.

As sex estimation techniques continue to evolve, practitioners will have to balance a range of factors such as time required for the analysis, time required for training, and accuracy of methods when choosing between traditional and computational approaches. This ongoing evaluation process will be crucial as researchers and practitioners work to refine and combine different methodologies, aiming to achieve the most accurate results in biological profiling across diverse populations. The needs of each technique will slightly differ depending on the overall goals of the analysis. For bioarchaeological research, an argument can be made that time can be allocated for a more computational approach. However, in relation to a forensic case, time is often of the essence; therefore, a straightforward, quickly, and accurate method is preferable.

In the context of bioarchaeology, the analysis of sexual variation within a sample population offers valuable insights into demographic structures, providing a nuanced understanding of past societies. This approach reveals intriguing patterns in classification scores across different populations (İşcan 2005). However, a key consideration when using GMM methods for bioarchaeological cases is preservation differences because whatever region is being focused on in GMM needs to be intact. This is more problematic in some areas of the skeleton compared to others (Waldron 1987; Stojanowski, Siedmann, and Doran 2002). The GSN is relatively robust and therefore is often used for GMM. For example, if all that is available is a fragmented *os coxa*, then utilizing GMM on the GSN will give the researchers a better sex estimation than using a purely visual technique (Conner et al. 2024). However, if focusing a more fragile aspect of the skeleton, such as the *os pubis*, visual methods may be superior as the inconsistency of preservation may exclude a large number of samples from GMM analysis.

Overall, the variability observed across populations (Gómez-Valdés et al. 2012; Velemínska et al. 2013; Kilmer and Garvin 2020; Conner et al. 2024) underscores the necessity for broader, more inclusive methodologies in sex estimation. While GMM techniques offer a potential higher precision, their application is often limited to specific populations, potentially overlooking the broader spectrum of human variation. In contrast, visual methods such as those presented here provide a more adaptable framework for assessing multiple populations, albeit with potential trade-offs in precision.

5 | Limitations

One of the biggest limitations of this study is that the sample population, albeit large (202 individuals), is only from one geographic location and time period. To fully understand the benefits and limitations of these methods, comparisons need to be made across geography and time to assess population specificity and secular changes. Everything possible was used to mitigate human error in digitizing the greater sciatic notches for the EFA; however, it cannot be fully excluded as a limitation.

Adapting the Walker (2005) classification system for the GSN may introduce a bias that negatively impacts male classifications by shifting the scoring system as seen in Table 1. As this adaptation was also used in previously published articles (Gómez-Valdés et al. 2012; Velemínska et al. 2013; Stinnesbeck et al. 2020; Conner et al. 2024; Russell et al. 2024), the lower accuracy for males using the Walker (2005) is more likely be due to population differences or secular changes, rather than using the adapted scoring system.

6 | Conclusion

Sex estimation is a key component of determining a biological profile. Although current methods have reasonable levels of accuracy, critiques often point to the subjectivity of the methods, especially when employed by researchers with limited experience. This study has shown that shape analysis yields comparable results to Bruzek's (2002) visual method on the GSN and some other digital methods (Kilmer and Garvin 2020, Conner et al. 2024) which would allow for a more objective determination. The authors suggest that further application of shape analysis and GMM using a larger and more diverse sample size would provide more insight into the application of these methods in bioarchaeology.

The ongoing debate between population-specific methods and broadly applicable methods highlights the complex nature of sex estimation in bioarchaeology and forensic anthropology. The potential to create more detailed and specific population profiles is dependent on improving the methods used in biological profiling. As the field continues to evolve, striking a balance between precision and applicability remains a key challenge within the forensic and archaeological community.

Acknowledgments

The authors would like to thank Margaret Clegg and Robert Kruzynski for access to the Spitalfields Coffin Plate Collection housed at the Natural History Museum, UK. Thanks to Matthew Hill and Stephen Axon for going over the readability of the article. We would also like to thank the anonymous reviewers for their comments which helped shape the article to be stronger.

Data Availability Statement

Data available on request from the authors. The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Acsádi, G., and J. Nemeskéri. 1970. *History of Human Life Span and Mortality*. Budapest: Akadémiai Kiadó.

Asala, S. A. 2001. "Sex Determination From the Head of the Femur of South African Whites and Blacks." *Forensic Science International* 117: 15–22.

Asala, S. A. 2002. "The Efficiency of the Demarking Pint of the Femoral Head as a Sex Determining Parameter." *Forensic Science International* 127: 114–118.

Bertsatos, A., M. E. Chovalopoulou, J. Brůžek, and Š. Bejdová. 2020. "Advanced Procedures for Skull Sex Estimation Using Sexually Dimorphic Morphometric Features." *International Journal of Legal Medicine* 134: 1927–1937. https://doi.org/10.1007/s00414-020-02334-9.

Bonhomme, V., S. Picq, C. Gaucherel, and J. Claude. 2014. "Momocs: Outline Analysis Using R." *Journal of Statistical Software* 56, no. 13: 1–24. https://doi.org/10.18637/jss.v056.i13.

Bruzek, J. 2002. "A Method for Visual Determination of sex, Using the Human Hip Bone." *American Journal of Physical Anthropology* 117, no. 2: 157–168. https://doi.org/10.1002/ajpa.10012.

Brůžek, J., F. Santos, B. Dutailly, P. Murail, and E. Cunha. 2017. "Validation and Reliability of the Sex Estimation of the Human Os Coxae Using Freely Available DSP2 Software for Bioarchaeology and Forensic Anthropology." *American Journal of Physical Anthropology* 164, no. 2: 440–449. https://doi.org/10.1002/ajpa.23282.

Buikstra, J. E., and D. H. Ubelaker. 1994. *Standards for Data Collection From Human Skeletal Remains: Proceedings of a Seminar at the Field Museum of Natural History, Organized by Jonathan Hass*. Fayetteville: Arkansas Archeological Survey.

Conner, L., A. Evin, L. Evis, C. McKenzie, K. Plomp, and C. Ameen. 2024. "Estimating Osteological Sex Using Predictive Geometric Morphometric Analysis of the Greater Sciatic Notch." *Journal of Archaeological Science: Reports* 58: 104745. https://doi.org/10.1016/j.jasrep.2024.104745.

Correia, H., S. Balseiro, and M. De Areia. 2005. "Sexual Dimorphism in the Human Pelvis: Testing a New Hypothesis." *Homo* 56, no. 2: 153–160. https://doi.org/10.1016/j.jchb.2005.05.003.

Cuzzulin, M. C., F. Curate, A. Rodrigues Freire, et al. 2020. "Validation of Anthropological Measures of the Human Femur for Sex Estimation in Brazilians." *Australian Journal of Forensic Sciences* 54, no. 1: 61–74. https://doi.org/10.1080/00450618.2020.1729411.

Franklin, D., A. Cardini, A. Flavel, and A. Kuliukas. 2012. "The Application of Traditional and Geometric Morphometric Analyses for Forensic Quantification of Sexual Dimorphism: Preliminary Investigations in a Western Australian Population." *International Journal of Legal Medicine* 126: 549–558. https://doi.org/10.1007/s00414-012-0684-8.

Gillet, C., L. Costa-Mendes, C. Rérolle, N. Telmon, D. Maret, and F. Savall. 2020. "Sex Estimation in the Cranium and Mandible: A Multislice Computed Tomography (MSCT) Study Using Anthropometric and Geometric Morphometry Methods." *International Journal of Legal Medicine* 134: 823–832. https://doi.org/10.1007/s00414-019-02203-0.

Gómez-Valdés, J. A., M. Quinto-Sánchez, A. M. Germandia, J. Velemínska, G. Sánchez-Mejorada, and J. Bruzek. 2012. "Comparison of Methods to Determine Sex by Evaluating the Greater Sciatic Notch: Visual, Angular and Geometric Morphometrics." *Forensic Science International* 221: 156.e1–156.e7. https://doi.org/10.1016/j.forsciint.2012.04.027.

Gunstra, N. D. S., L. Betti, B. Fischer, et al. 2023. "There Is an Obstetrical Dilemma: Misconceptions About the Evolution of Human Childbirth and Pelvic Form." *American Journal of Biological Anthropology* 181, no. 4: 535–544. https://doi.org/10.1002/ajpa.24802.

Harma, A., and H. M. Karakas. 2007. "Determination of Sex From the Femur in Anatolian Caucasians: A Digital Radiological Study." *Journal of Forensic and Legal Medicine* 14: 190–194.

Herzlinger, G., N. Goren-Inbar, and L. Grosman. 2017. "A New Method for 3D Geometric Morphometric Shape Analysis: The Case Study of Handaxe Knapping Skill." *Journal of Archaeology Science: Reports* 14: 163–173. https://doi.org/10.1016/j.jasrep.2017.05.013.

Hoggard, C. S. 2020. "Geometric Morphometrics (GMM) and Archaeology Workshop (July 2020) [Software]." Github. https://github.com/CSHoggard/-gmm_liverpool_2020.

İşcan, M. Y. 2005. "Forensic Anthropology of Sex and Body Size." *Forensic Science International* 147, no. 2–3: 107–112. https://doi.org/10.1016/j.forsciint.2004.09.069.

Jovanonic, S., and S. Zivanovic. 1965. "The Establishment of the Sex by the Great Sciatic Notch." *Acta Anatomica (Basel)* 61: 101–107.

Kanahashi, T., J. Matsubayashi, H. Imai, S. Yamada, H. Otani, and T. Takakuwa. 2024. "Sexual Dimorphism of the Human Fetal Pelvis Exists at the Onset of Primary Ossification." *Communications Biology* 7, no. 1: 538. https://doi.org/10.1038/s42003-024-06156-y.

Kilmer, K., and H. Garvin. 2020. "Outline Analysis of Sex and Population Variation in Greater Sciatic Notch and Obturator Foramen Morphology With Implications for Sex Estimation." *Forensic Science International* 314: 1–9. https://doi.org/10.1016/j.forsciint.2020.110346.

Klales, A. R. 2016. "Secular Change in Morphological Pelvic Traits Used for Sex Estimation." *Journal of Forensic Sciences* 61: 295–301. https://doi.org/10.1111/1556-4029.13008.

Klales, A. R. 2020. *Sex Estimation of the Human Skeleton: History, Methods, and Emerging Techniques*. London: Academic Press.

Klales, A. R., S. D. Ousley, and J. M. Vollner. 2012. "A Revised Method of Sexing the Human Innominate Using Phenice's Nonmetric Traits and

Statistical Methods." *American Journal of Physical Anthropology* 149, no. 1: 104–114. https://doi.org/10.1002/ajpa.2210.

Klingenberg, C. P. 2011. "MorphoJ: An Integrated Software Package for Geometric Morphometrics." *Molecular Ecology Resources* 11: 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x.

Kranioti, E. F., N. Vorniotakis, C. Galiatsou, M. Y. Iscan, and M. Michalodimitrakas. 2009. "Sex Identification and Software Development Using Digital Femoral Head Radiographs." *Forensic Science International* 189: 113.e1–113.e7.

Kurki, H. K. 2007. "Protection of Obstetric Dimensions in a Small-Bodied Human Sample." *American Journal of Physical Anthropology* 133: 1152–1165. https://doi.org/10.1002/ajpa.20636.

Kurki, H. K. 2011. "Pelvic Dimorphism in Relation to Body Size and Body Size Dimorphism in Humans." *Journal of Human Evolution* 61: 631–643. https://doi.org/10.1016/j.jhevol.2011.07.006.

Kurki, H. K. 2013a. "Bony Pelvic Canal Size and Shape in Relation to Body Proportionality in Humans." *American Journal of Physical Anthropology* 151, no. 1: 88–101. https://doi.org/10.1002/ajpa.22243.

Kurki, H. K. 2013b. "Skeletal Variability in the Pelvis and Limb Skeleton of Humans: Does Stabilizing Selection Limit Female Pelvic Variation?" *American Journal of Human Biology* 25: 795–802. https://doi.org/10.1002/ajhb.22455.

Landis, J. R., and G. G. Koch. 1977. "An Application of Hierarchical Kappa-Type Statistics in the Assessment of Majority Agreement Among Multiple Observers." *Biometrics* 33: 363–374. https://doi.org/10.2307/2529786.

Lazorthes, G., and A. Lhez. 1940. "Les caractés sexuels de la grande échancrure sciatiqu." *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 10: 180–185.

Liutu, T., and P. M. Dixon. 2020. "The use of the Geometric Morphometric Method to Illustrate Shape Difference in the Skulls of Different-Aged Horses." *Veterinary Research Communications* 44: 137–145. https://doi.org/10.1007/s11259-020-09779-8.

Lovell, N. C. 1989. "Test of Phenice's Technique for Determining Sex From the *Os Pubis*." *American Journal of Physical Anthropology* 79, no. 1: 117–120. https://doi.org/10.1002/ajpa.1330790112.

MacLaughlin, S., and M. Bruce. 1990. "The Accuracy of Sex Identification in European Skeletal Remains Using the Phenice Characters." *Journal of Forensic Sciences* 35, no. 6: 1384–1392. https://doi.org/10.1520/jfs12974j.

Maio, C., E. Cunha, and D. Navega. 2024. "Metric Analysis of the Patella for Sex Estimation in a Portuguese Sample." *Forensic Sciences Research* 9, no. 2: owae015. https://doi.org/10.1093/fsr/owae015.

Molleson, T., and M. Cox. 1993. "The Spitalfields Project, Vol. 2: The Anthropology. The Middling Sort." In *Council for British Archaeology Research Report 86*. York: Council for British Archaeology.

Murphy, A. M. C. 2005. "The Femoral Head: Sex Assessment of Prehistoric New Zealand Polynesian Skeletal Remains." *Forensic Science International* 154: 210–213.

Phenice, T. W. 1969. "A Newly Developed Visual Method of Sexing the Os Pubis." *American Journal of Physical Anthropology* 30: 297–302.

Reeves, J., and M. Adams. 1993. "The Spitalfields Project, Vol. 1: The Archaeology. Across the Styx." In *Council for British Archaeology Research Report 85*. York: Council for British Archaeology.

Rennie, S. R. 2018. "Summary Sex: A Multivariate Approach to Sex Estimation From the Pelvis." PhD Thesis, Liverpool John Moores University, UK.

Rennie, S. R., C. Eliopoulos, and S. Gonzalez. 2023. "Evaluation of the Obturator Foramen as a Sex Assessment Trait." *Forensic Science Medicine and Pathology* 19: 146–153. https://doi.org/10.1007/s12024-022-00514-0.

Russell, M., M. Smith, E. Hambleton, P. Cheetham, and H. Tamminen. 2024. "Brutalised, Bound and Bled: A Case of Later Iron Age Human Sacrifice From Winterbourne Kingston, Dorset." *Antiquaries Journal*: 1–27. https://doi.org/10.1017/S0003581524000143.

Singh, S., and B. R. Potturi. 1978. "Greater Sciatic Notch in Sex Determination." *Journal of Anatomy* 125: 619–624.

Spradley, M. K., R. L. Jantz, A. Robinson, and F. Peccerelli. 2008. "Demographic Change and Forensic Identification: Problems in Metric Identification of Hispanic Skeletons." *Journal of Forensic Sciences* 53: 21–28. https://doi.org/10.1111/j.1556-4029.2007.00614.x.

Stinnesbeck, W., S. R. Rennie, J. Avilés Olguín, et al. 2020. "New Evidence for an Early Settlement of the Yucatán Peninsula, Mexico: The Chan Hol 3 Woman and Her Meaning for the Peopling of the Americas." *PLoS ONE* 15, no. 2: e0227984. https://doi.org/10.1371/journal.pone.0227984.

Stojanowski, C. M., R. M. Siedmann, and G. H. Doran. 2002. "Differential Skeletal Preservation at Windover Pond: Causes and Consequences." *American Journal of Physical Anthropology* 119: 15–26. https://doi.org/10.1002/ajpa.10101.

Velemínska, J., V. Krajíček, J. Dupej, et al. 2013. "Technical Note: Geometric Morphometrics and Sexual Dimorphism of the Greater Sciatic Notch in Adults From Two Skeletal Collections: The Accuracy and Reliability of Sex Classification." *American Journal of Physical Anthropology* 152: 558–565. https://doi.org/10.1002/ajpa.22373.

Waldron, T. 1987. "The Relative Survival of the Human Skeleton: Implications for Palaeopathology." In *Death, Decay and Reconstruction: Approaches to Archaeology and Forensic Science*, edited by A. Boddington, A. N. Garland, and R. C. Janaway, 55–64. Manchester: Manchester University Press.

Walker, P. L. 2005. "Greater Sciatic Notch Morphology: Sex, Age, and Population Differences." *American Journal of Physical Anthropology* 127: 385–391. https://doi.org/10.1002/ajpa.10422.

Walker, P. L. 2008. "Sexing Skulls Using Discriminant Function Analysis of Visually Assessed Traits." *American Journal of Physical Anthropology* 136, no. 1: 39–50. https://doi.org/10.1002/ajpa.20776.

Workshop of European Anthropologists. 1980. "Recommendations for Age and Sex Diagnoses of Skeletons." *Journal of Human Evolution* 9: 517–549.