Universitat Rovira i Virgili
Departament d'Història i Història de l'Art
Màster en Arqueologia del Quaternari i Evolució Humana

International Master in
QUATERNARY AND PREHISTORY

Master's Thesis

New Methodological Advances in the Study of Taphonomic Equifinality in the Lower Pleistocene Site of FLK-West (Olduvai Gorge, Tanzania)

Lloyd Austin David Courtenay

Directors: Rosa Huguet and José Yravedra

Curso académico 2017/2019

Aknowledgements

Trying to think of who to thank for the last two years while preparing and writing my Master's Thesis seems almost impossible to write, let alone summarise.

To try and begin, first I would like to thank absolutely all the members and staff of the Rovira I Virgili University as well as the IPHES lab. I would like to express my gratitude especially to the teachers of this Master's Degree who have all contributed in my academic formation over the last two years. I would like to especially thank Andreu Ollé, Josep M. Vergès and Marina Mosquera for their passion and inspiration; teaching me a great deal about subjects I knew little about before and have now come to enjoy greatly. I would also like to thank Carlos Lorenzo for the advice he gave on a number of occasions. Additionally, I would like to express my gratitude to the Anatomie Comparée of the Muséum National d'Histoire Naturelle for providing access to the African Fauna collections. From the Musée del Homme I would additionally like to thank Florent Detroit and Martin Friess for their interesting lectures and insights on Geometric Morphometrics. Along these lines I would also like to thank my colleagues and companions who have shared this learning experience with me in class. This is not exclusive to my time here in Tarragona, but also extends out to all the people I met during my stay in France, especially Marie, Alexis, Pierre, and I would especially like to thank Julia for her enthusiasm and obsession with trying to get me to like Paris.

The list of friends I have made over these last years in Spain and those who have supported me is too long to name individually. As a summary to these particularly interesting individuals I have this to say: you know who you are, and thank you for being there.

To provide some honourable mentions, first I would like to thank all the friends I made during my degree at the Complutense, while we have recently parted ways I hope to see you all soon. Alongside this, I would like to thank all the people I have met over the years in the various countries digging. I think an archaeological dig is the best place to meet some of the most colourful and unusual people ever. Digging is the most enriching experience from a personal as well as professional level, and I have enjoyed each one a great deal. With this I would like to thank especially the team members I have met while digging.

From the IPHES there are many, many people I would like to thank, but it would be impossible to name you all individually. A special thanks, however, has to go to all of those who would end up in the Metal Heart most weekends until they kicked us out… which reminds me, I would also like a very special acknowledgement to go out to *Beer*. Without it, this Master's Thesis would have been impossible.

Moving on to include the list of people who have helped in my formation as a professional, the first people I would like to acknowledge are those who have helped considerably in my formation in taphonomy. With this I would like to give my thanks to the minds of Julia Aramendi, Lucia Cobo-Sánchez, Pablo Lopez Cisneros, Julio Rojo Hernández, Cristina Gómez-Miguelsanz, and most recently Antonio Rodríguez-Hidalgo, Isabel Cáceres, Jordi Rosell, Palmira Saladié and Antonio Pineda. Finally, I would like to express my utmost gratitude to Abel Moclán for teaching, helping and giving me advice on many occasions. As Abel is probably aware, I truly appreciate and value his opinion greatly.

Moving on, I would like to give a special thanks to Julia Aramendi and Miguel-Ángel Maté-González who I have had the pleasure to work with over the last couple of years. I value all the effort that Julia has given me to improve my work as an author and all she has taught me about geometric morphometrics. I also would like to thank Miguel-Ángel for all the support, motivation, attention and help he has provided over the years. I am very grateful to have worked with both since beginning in this field. I would also like to thank Diego González-Aguilera and the TIDOP lab in Avila for their support in the investigation we do. I would finally like to thank Andreu Ollé again for all his support, help and enthusiasm.

Next I would like to express my gratitude to Manuel Domínguez-Rodrigo, my teacher and a considerable inspiration for most of my work. I would like to thank him for his innovative research and for introducing me to the world of taphonomy. Alongside this, on a professional level, I would also like to thank him for providing access to the experimental materials used to develop a large part of this study.

Moving on I would like to thank the Olduvai Team, especially those who made my experience in TK so enjoyable. I would like to also thank all those working at FLK-West, especially Fernando Diéz-Martín. I would like to thank absolutely everyone I met in Tanzania, including all the Tanzanians, Masai and the country itself for just being incredible. If it's possible to thank the African continent, I would do that too. I would like to express my gratitude to Joaquín Panera and Susana Rubio-Jara for introducing me to the country and making me feel so welcome, especially the first couple of days in Arusha. I can't express how happy the whole experience of the Olduvai Gorge made me feel, so a thanks goes out to everyone who was involved in such a life-changing experience.

The following acknowledgements go out to those closest to me, starting with John Banks for motivating and supporting everything I do. I am incredibly pleased to include you in my family and appreciate all you do for me.

Next, my closest friends. First, my friend Juan. Because you can hardly speak a word of English, I provide the following in Spanish, just for you: "*Gracias Juan por toda la musica y, sobre todo, el lúpulo*." With Juan I send out my thoughts to all those in La Tienda de la Cerveza, no stop in Madrid is complete without it. Next I thank Enrique, for always being there. **Always**. Finally, to the three people who have had to put up with me for these last two years and who have turned in to some of my closest friends; Andrea, Miguel and Noé. I can't express in the words of any language how much I appreciate you all.

My highest level of gratitude goes towards my two directors. I would like to thank Rosa Huguet for everything she has done for me over the past two years. All the advice, help and lessons… I could never thank her enough. I feel that as a professional I have grown greatly with her guidance, and would like to acknowledge her patience especially. I know I'm not the easiest person to put up with and would like to express how much I appreciate the contributions she has made to my work and my formation as a professional. Finally I would like to acknowledge and show my gratitude to one of my greatest academic inspirations and mentor, José Yravedra, without whom, none of this would have been possible.

Last but not least, I want to express my admiration for the entire Courtenay family whose name I bear so proudly. With this, the final people I would like to thank cannot be expressed as an acknowledgement. Instead:

All my work is dedicated to:

The person who I miss more than anything

Ginger Courtenay

The person who is always there for me

Liz Courtenay

The person who always understands me

Jordan Courtenay

Abstract

The concept of equifinality is currently one of the largest issues in taphonomy, frequently leading analysts to erroneously interpret the formation and functionality of archaeological and paleontological sites. This issue lies primarily in the methods available to process microscopic bone surface modifications, especially when considering superficial traces such as naturally produced trampling marks when compared with cut marks. Here the 1.7 Ma Acheulean site of Frida Leakey Korongo West (Olduvai Gorge, Tanzania) is used as a testing ground for new methodological approaches designed to confront these issues. This study combines advanced digital microscopy for the study of superficial taphonomic traces, geometric morphometrics and other statistical techniques, such as Machine *and* Deep Learning*, to process problematic taphonomic data. As a result, this Master's Thesis presents the first archaeological implementation of Artificial Intelligence in the processing of morphological data extracted using high resolution digital microscopy, obtaining up to 100% classification in some cases of taphonomic alterations. These techniques can thus be considered a promising tool for the future of Archaeological Science.*

Key Words: Taphonomy, Artificial Intelligence, Geometric Morphometrics, Microscopy, Olduvai Gorge.

Résumé

Le concept de l'équifinalité est actuellement l'un des plus grands problèmes dans la taphonomie, ce qui cause souvent l'interprétation erronée de la formation et de la fonctionnalité des sites archéologiques et paléontologiques. Ce problème est dû principalement aux méthodes disponibles de développer les modifications sur les surfaces osseuses microscopiques, particulièrement considérant les traces superficielles comme les marques de piétinement naturellement produites comparées aux marques de découpe. Dans ce travail le site acheuléen de Frida Leakey Korongo West (la gorge d'Olduvai, Tanzanie), daté d'il y a 1.7 millions d'années, est utilisé comme un terrain d'essai pour les nouvelles approches méthodologiques qui sont conçues pour affronter cette problématique. Cette étude combine la morphométrie géométrique avec des autres techniques statistiques avancées, comme le Machine *et* Deep Learning*, afin de traiter les données taphonomiques problématiques. Les résultats de ce mémoire de maitrise présentent la première utilisation archéologique de l'Intelligence Artificielle dans le traitement d'information morphologique, arrivant à obtenir jusqu'à 100% de succès dans la classification dans certains cas de quelques modifications taphonomiques. Donc ces techniques peuvent être considérées comme un outil promettant pour l'avenir de la Science Archéologique.*

Mots Clés: Taphonomie, Intelligence Artificielle, Morphométrie Géométrique, Microscopie, la Gorge d'Olduvai.

Resumen

El concepto de la equifinalidad se puede considerar uno de los problemas actuales más grandes en la tafonomía, siendo la causa más frecuente de la interpretación errónea de la formación y funcionalidad de los yacimientos arqueológicos y paleontológicos. Dicho problema se debe principalmente a los métodos disponibles actualmente para procesar las alteraciones microscópicas sobre las superficies óseas. Un ejemplo de esto se encuentra en la diferenciación entre las marcas de cortes antrópicas y las marcas producidas por agentes naturales como como es el ejemplo del pisoteo. Este trabajo presenta el yacimiento Achelense Frida Leakey Korongo West (la Garganta de Olduvai, Tanzania), datado de hace 1,7 Ma, como un caso de estudio práctico para la aplicación de nuevas metodologías diseñados para enfrentar dicha problemática. Este estudio combina la microscopía digital avanzada para el estudio de marcas superficiales, la morfometría geométrica con otros métodos estadísticos avanzados, incluyendo el Machine *y* Deep Learning*, para poder procesar datos tafonómicos problemáticos. Los resultados de este Trabajo Fin de Máster presentan por primera vez el uso de Inteligencia Artificial para procesar datos morfológicos obtenidos utilizando la microscopia digital de alta resolución en el campo de la arqueología, llegando a obtener hasta el 100% de éxito en la clasificación en algunos casos de alteraciones tafonómicas. Estas técnicas se pueden considerar como una herramienta potente para el futuro de la Ciencia Arqueológica.*

Palabras Claves: Tafonomía, Inteligencia Artificial, Morfometría Geométrica, Microscopía, Garganta de Olduvai.

New Methodological Advances in the Study of Taphonomic Equifinality in the Lower Pleistocene Site of FLK-West (Olduvai Gorge, Tanzania)

INDEX

Figure Index

Table Index

Table S14 – (Page 258) – Initial classification of carnivore taxa using tooth pits from FLK-West using all 3 ML models and the 1 DL model.

East Africa is a landscape in turmoil. Torn apart, by the twisting and buckling of the Earth's crust. East Africa is also a landscape of huge and unpredictable change. The animals here are forced, day by day, season by season, to gamble with their lives. But for those that win, this is one of the most fertile landscapes on Earth.

…

East Africa may sometimes appear cruel, but there is nowhere else that provides such rich opportunities for those that are prepared to take them. And it was these ever-changing Savannahs that produced the most adaptable species of all:

Ourselves

Sir David Attenborough

Africa

Episode 2 - Savannah

Chapter 1

Introduction

1.1. Presentation

The Master's Thesis titled "*New Methodological Advances in the Study of Taphonomic Equifinality in the Lower Pleistocene Site of FLK-West (Olduvai Gorge, Tanzania)"* presents almost two years of research, with a background that stretches back to the early stages of the 2010's with earlier works directed by José Yravedra, and further developed through observations produced in my Dissertation (Courtenay, 2017). This Master's Thesis can thus be considered an expansion of previous efforts in taphonomic analysis using new technologies which were originally able to reveal a new means of processing cut mark morphology with significant differences between samples (research begun by Maté-González et al., 2015 and later developed and published officially in Courtenay et al., 2017). Nevertheless, these applications to the archaeological register began to produce mixed results, with significant developments in the case of the Oldowan site of Bell's Korongo (BK), Olduvai Gorge, Tanzania, (Courtenay et al., Under Review), but disappointing results for the case of Frida Leakey Korongo West (FLK-West).

With this, this present body of work intends to confront these issues, searching for new means of processing taphonomic information using a multitude of different approaches. In order to present these research efforts, this study has been structured in a number of different sections and subsections that will be described in continuation.

Chapter 2 of this study includes a contextualisation of all the efforts and methods that have been produced over the last century in taphonomic research, with an introductory theoretical presentation of new techniques that are about to be introduced into archaeological research. These consist in the presentation of taphonomic theory, followed by more precise concepts within this field. Among these sections, the *Theoretical Framework* will present some key components of scientific research within archaeology, presenting the different analytical methods available to study the fossil record. These include experimentation following scientific protocol, microscopy for in depth analysis of bone surfaces, geometric morphometrics for the quanitification of morphologies, and finally, the definition and description of a new scientific field that has been recently introduced into archaeology; Artificial Intelligence.

Following this, Chapter 3 describes the case study at hand where the different methods will be evaluated. Here the historiography of the Olduvai Gorge is explained in detail, with a brief presentation of the most important research that has been performed in this important component of Eastern African heritage. Finally, the precise case of FLK-West is presented, highlighting the preliminary finds uncovered from this important site and the impact their study may have on our understanding of early hominin populations in the Lower Pleistocene.

Chapter 4 is dedicated to the samples and methods used to carry out this research, providing in depth descriptions of the experimental material used and how they have been studied. In this section, a presentation will be provided for each of the pieces of equipment used, such as the HIROX KH-8700 3D Digital Microscope, followed by explanations of the different components of data science and artificial intelligence applied in the processing of the data produced, as described in Chapter 5.

Chapter 6 confronts the contextualisation of the data produced, and the impact these methods may have on future research efforts into archaeology as well as palaeontology.

Next, Chapter 7 is divided into 2 parts, highlighting not only the conclusions withdrawn from the presented research, but also the proposals for future investigations.

The 8th part of this Master's Thesis includes the full list of bibliographical references used to construct this study. Here over 800 references to research efforts in

various different fields, including both archaeology and artificial intelligence, are used to support the data provided and construct the arguments presented.

Finally, this is followed by a series of Appendices, providing a glossary of definitions for key concepts, a list of abbreviations used throughout the study and two Appendices including additional graphs and photographs obtained throughout the course of the last two years of work. The final two appendices are dedicated to the presentation of two articles published as product of this Master's Thesis, both having been peer reviewed and published in high impact journals within science and archaeology.

Before any of this can begin, however, the remainder of this small introductory chapter is dedicated to the presentation of the null hypotheses and objectives that have inspired this body of research.

1.2. Objectives

Bone Surface Modifications (BSMs) have been an important protagonist in taphonomic investigation for many years, especially in light of the confusion created when trying to classify the agent responsible for producing them. The phenomenon that describes our ability to misinterpret a taphonomic trace can be summarised by the term *Equifinality*.

Equifinality has been the cause for the erroneous interpretation of some archaeological and paleontological sites for decades. The similarity between traces produced by many agents have even reached the point where our understanding of site formation can change drastically when considering the analyst reviewing this data. In some cases paleontological sites have even become classed as archaeological, leading to remarkable claims about our understanding of human evolution.

In many cases zooarchaeological analysis can be performed without the need for overly complex pieces of equipment. The classification of taxa, anatomical elements, mortality patterns, pathologies, and other features can at times be identified easily using the naked-eye, requiring experience and possible reference collections for comparison. From another perspective taphonomic analyses of bone surface modifications can be performed using simple 10 to 40 x magnification eye-lenses. Not all studies, however, can be efficiently performed in this manner, and may require high resolution techniques in order to confront certain investigation questions, such as those observed in the site of Frida Leakey Korongo West (FLK-West, Olduvai Gorge, Tanzania). Here lies the inspiration for the present Master's Thesis.

Recent studies confronting the first faunal remains to be uncovered from FLK-West by Yravedra et al. (2017a) and Yravedra and Domínguez-Rodrigo (2018) have noticed significant evidence of anthropic butchery activities, especially in the lower levels of this site. Nevertheless, these authors have also noted a significant presence of fluvial abrasion and polishing to cortical surfaces, thus effecting the BSMs observed. Said alterations, have produced significant difficulties when processing these traces and in withdrawing conclusions as to their classification (Yravedra et al., 2017a; Courtenay, 2017).

In recent years, a major advance that has begun to confront these issues is found in the development and introduction of new technologies into the field of taphonomy. From this perspective we find the use of geometric morphometrics, advanced statistics and the digital reconstruction of traces for more in depth analysis of their formation.

The objectives of this body of research, therefore, are twofold; on one hand this study intends to make an attempt at developing new

technologies that can be used to remove problems of equifinality in the taphonomic register. Secondly, the consequent research intends to use the preliminary samples obtained from the African Lower Pleistocene Acheulean site of FLK-West as a testing ground for these methodological developments.

A possible solution in overcoming issues of equifinality relies on the use of microscopy for the study of superficial traces. Here, due to availability of different tools for taphonomic analysis at the laboratories where this study has been performed, the HIROX KH-8700 3D Digital Microscope is used as the primary piece of equipment. Nevertheless, the HIROX employs light to reconstruct the subject under observation, presenting the problem of how this variable may affect the quality of the reconstruction based on its position over the object. Therefore, the evaluation of the HIROX as a tool for taphonomic studies, comparing this microscope with other techniques that have been used for micromorphological analyses.

Once the optimum protocol for the microscope has been defined, the next objective is to test out its performance when confronted with superficial traces such as trampling marks. Upon analysing these types of microscopic BSMs, a characterisation of trampling marks and their comparison with other traces such as cut marks can be performed.

An additional line of research that became increasingly apparent throughout the statistical processing of data lay in the selection of statistical tests being used. Upon in depth bibliographical research and statistical experimentation, a series of algorithms with great potential for studying this type of data were prepared. This additional objective meant that these new techniques had to be evaluated in comparison with older methods, finding out which of the algorithms were the most efficient and robust for the study of archaeological databases.

Once the methodology is defined and evaluated, the final phase of this Master's Thesis is to confront the second main objective described before; testing out the different approaches on the archaeological record.

The development and use of new technologies and statistical models in archaeology can be considered an important component of evolving as a discipline. This Master's Thesis intends to prove how an updated use of new microscopic and statistical techniques can be used to process problematic archaeological data. While these conclusions are still unable to overcome *all* taphonomic problems, the underlying intentions are to present a new means of investigation, opening multiple doors to the future of archaeological science in the study of Human Evolution.

Chapter 2

Theoretical Framework

2.1. The Roots of Taphonomy

The coining of the term *taphonomy* was originally conceived in the academic field of Paleontology; what this term intended to describe were the series of processes that, as a whole, form a paleontological excavation site. The roots of this term come from the Greek words for 'burial' (*taphos*) and 'laws' (*nomos*), creating the rather generic term; site formation processes. The first use of the word *taphonomy* can be found in the work of Russian palaeontologist Ivan Efremov (1940), whose intentions lay in finding a new methodology that could aid in the study of paleontological remains. The reference to these "laws of embedding" were Efremov's way of recovering all the possible information behind the formation of a site. Since the conception of this term, however, the definition of *taphonomy* has undergone a series of long complex processes that essentially dissected and developed this term, producing the inevitable redefinition of the word *taphonomy* as well as the academic field.

Within the paleontological record, taphonomy began purely as the biological study of a site with its key objectives rooting in the understanding of paleobiological evolutions. According to Efremov, these paleobiological evolutions begin with the death of an organism, followed by a sedimentary process that eventually buries the remains of said organism (*biostratinomy*). Taphonomic processes continue to act upon these remains (*diagenesis*) until a final phase where the remains are eventually discovered and analysed. In this sense, taphonomy tries to deduce the series of processes that could have affected the preservation of the paleontological record, taking into account the numerous taphonomical '*agents'* that could have altered the preservation of these organic remains. From this we can extract the logical assumption that over the passing of time only a small amount of the original organic material have received a level of preservation that has contributed to their survival (Darwin, 1859; Montelius, 1888; Efremov, 1940). In order to understand why such a small percentage of the original remains have survived, however, we have to understand all the processes that contribute to the formation of the fossil record.

Palaeontologists began developing the concept of taphonomy stating the need to strip away of the "taphonomic overprint", and thus "decipher the paleocological significance displayed" in a site (Lawrence, 1979; Wilson, 1988). From this, the term *taphonomy* received a new definition in which taphonomy can be defined as the reconstruction and formation of the fossil record in order to understand the true nature behind its accumulation (Behrensmeyer et al., 2000). The eventual evolution of taphonomy in this manner presented the gradual application of this field towards archaeological remains. This change can be relatively dated back to the 1970's, through references towards "non-cultural formation process concepts" in the explanation of the interactions between archaeological remains and the environmental conditions in which they were deposited (White, 1979). The problem with White's initial integration of paleontological and archaeological sites, however, exists in the words "non-cultural", as it excludes any modification to organic remains that could have been presented through anthropic interactions with said organism (Lyman, 2010).

An important revolution within the formation of the taphonomic discipline, especially with reference to archaeological sites, was the definition of taphonomy as a study that not only includes organic remains but also takes into account non-organic remains that form part of paleontological/archaeological sites. An example of this can be seen through the

inclusion of lithic tools, as well as the faunal remains that are associated with them. This distinction becomes highly important when including taphonomical studies in the archaeological record, taking into account that the best way of studying the past lies within the integration of various academic fields (Binford, 1962; Clarke, 1973; Renfrew, 2014). While taphonomy had previously been classed as an almost purely organic and biological field, ignoring the presence of non-organic elements that also form an important part of any archaeological site; the importance of these non-organic elements also provide information that can be used to deduce the original environment in which certain remains have been deposited. As pointed out by some critics (Dominguez-Rodrigo et al., 2011a), Efremov, in his first definition of taphonomy (Efremov, 1940), never limited the study to purely organic material. Misunderstandings of Efremov's work, however, have led analysts to concentrate their studies solely on organic materials, most specifically referring to faunal remains. Efremov's focus on fauna, however, can clearly be seen by his dedication to the paleontological field and its study of "terrestrial animal remains".

Theoretically, non-organic remains can also contribute information on the contextual positioning of osteological finds than the bones themselves (seeing as lithic materials, for example, have higher preservation rates than organic material). In light of this, archaeological taphonomy should take into consideration the entire context of organic materials, whereby truly understanding the true nature of hominin behaviour and interactions with zooarchaeological finds. These non-organic materials could consist of lithic materials, raw materials, sedimentation processes and even the negative imprints of organic structures (Dominguez-Rodrigo et al., 2011a).

This particular theoretical observation has had a particularly powerful presence in other fields such as traceology (Semenov, 1964; Keeley, 1980; Xhauflair et al., 2017), residue analysis (Fullagar et al., 1996; Haslam, 2006), spatial archaeology (Marean and Bertino, 1994: Araujo and Marcelino, 2003; Camarós et al., 2013; Arilla et al., 2014; Domínguez-Rodrigo et al., 2017a).

In this evolution of the taphonomic academic field, a need to define what can be considered as a "taphonomic agent", as well as a "taphonomic process", has become essential in its integration and use in archaeological investigation. The formation of organic accumulations can currently be divided into 4 main phases; *biocoenosis*, *thanatocoenosis*, *taphocoenosis*, *oryctocoenosis*. As explained by Fernández López (2000), these phases can be defined as followed:

- *Biocoenosis* – A group or set of organisms that live in the same locality or region.
- *Thanatocoenosis* – A group or set of organic remains that died together.
- *Taphocoenosis* – A group or set of organic remains that have been buried together.
- *Oryctocoenosis* – A group or set of fossils that have been discovered together.

Among these phases, the concepts of biostratinomy and diagenesis are still valid, defining the modifications to have occurred before burial as well as any processes to have occurred after sedimentation has taken place (respectively). Fernández López continues to point out the necessity of other terminology such as thanatotope (geographical area occupied by a group of carcasses from a determinable class) and taphotope (geographical area where said class of carcasses are buried) that aid in the spatial understanding of site accumulations (Fernández López, 1991, 2000). The need

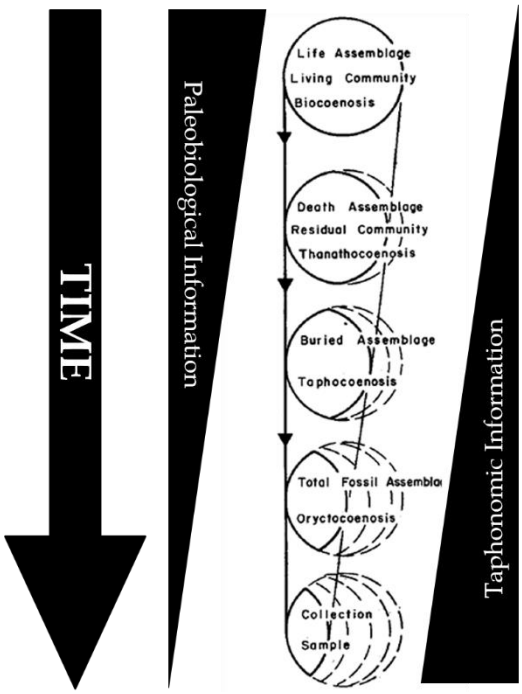


Figure 2.1 - Diagram explaining the loss and gain of information over the passing of time in the fossil and taphonomic record according to Fernández López (1991)

for these definitions is essential in our capacity to interpret fossil association, better defined as the "fossil-assemblage". This can be clarified over the course of time; while the paleobiological register begins to diminish, the taphonomical regiter is in a constant state of growth (Fernández López, 1991, Fig. 1.1.).

If we are to immerse ourselves within the archaeological field in relation to taphonomical information, we have to take into consideration our most important definition of archaeology; the understanding of human development and the (pre)history of our species (whether we take into consideration either biological evolution, or cultural/cognitive factors). As defined previously, taphonomy has to include natural as well as non-natural processes. If we were to expand on this we would also have to include cultural processes as well. In the end, all hominid activity is conditioned by a certain series of cognitive elements that can be interpreted and defined by the socio-cultural environment in which our species can be located. Taking this highly complex conditioning factor into account, humans can be defined as "just another taphonomic agent" (Dominguez-Rodrigo et al., 2011a). Part of archaeology's contribution to evolutionary studies has tried to approach the problematic question of "what makes us human?"; this question is essentially conditioned by the evidence at hand. If we are able to understand the taphonomical processes that affect the accumulation of said "evidence", we can then generate questions on how to reconstruct human behaviour and answer this question (among many others).

In understanding hominid behaviour we are confronted with a number of problems, the very problems described by Darwin (1859) and Efremov (1940) in their first reflections on the fossil record: the lack of quality in fossil conservation. In the archaeological record, what we are left with are simple "footprints" that these individuals have left behind. If we are to understand the *biocenosis* and interaction between hominid populations and the fauna that share an ecosystem, we have to look for these multiple "traces" that have been left behind (Gifford-Gonzalez, 1991). The activity that appears to be most important in archaeological sites, thus, becomes ancient "butchery" and the long series of debates that appear intertwined with this very subject. As Binford states in his book "Bones" (1981); "*when man butchers an animal, he leaves traces of his cutting and chopping activities*", a paradigm upon which many taphonomers, as well as (zoo)archaeologists, have concentrated their efforts in understanding. With the development of quantitative (Uerpmann, 1973; Binford, 1981; Grayson, 1984; Marshall and Pilgram, 1993; Lyman, 1994) and qualitative (Shipman and Rose, 1983; Olsen and Shipman, 1988; Capaldo and Blumenschine, 1994) analyses of faunal assemblages, new studies of African sites began debating the true hunting nature of early Hominin populations (Binford, 1981;

Bunn, 1981, 1982a; Binford, 1988a), eventually developing into what is now known as the "Hunter-Scavenger Debate" (Binford, 1981, 1983, 1984a, b, 1985, 1987, 1988a, b, c, 1989; Binford and Stone, 1987a, b against Bunn, 1981, 1982a, b, 1983a, b, 1986, 1987, 1989, 1993, 1994; Bunn and Kroll, 1986, 1987, 1988). This debate primarily revolved around the access of hominins to carcasses (Isaac, 1978, 1983; Blumenschine, 1986; Bunn and Edzo, 1993; Bunn, 1982a, 1993, 1994 1995, 1996), yet led to multiple discussions about the interpretation and classification of the taphonomic variables at hand (Blumenschine, 1988; Marean et al., 1992; Blumenschine and Marean, 1993; Selvaggio, 1994a, b; Blumenschine, 1995; Capaldo, 1995, 1997).

What begins to become clear within the evolution of this academic field is a chain of concepts resumed nicely in the work of Gifford-Gonzalez (1991). Key factors in taphonomy that become essential in the objectives behind this study are the concepts of *Trace*, *Casual Agent*, *Effector*, *Actor*, *Behavioural* and *Egological Contexts*. The *Trace* is the mark observed by the analyst, the *Casual Agent* being the process that leads to this *Trace*'s production. The *Effector* is a form of intermediate element between the *Actor* (the individual or taphonomic agent producing the mark) and the *Casual Agent*. *Behavioural* and *Ecological Contexts* are rather self-explanatory, however, these two concepts at times become the more interesting parts of the chain and, to our disadvantage, some of the harder concepts to interpret and imply. To provide a simple example of this chain we can try and imagine a group of Hominins (*Homo ergaster*, for example) as our *Agent*, in their ecological niche (the savannahs of Africa). In their necessity to eat, a group of *Homo ergaster* hunt a wilder beast (the *Behavioural Context*) and begin by processing the cadaver (*Casual Agency*). Using a simple flake (the *Effector*) made from local raw materials (also forming part of the *Ecological Context*) they begin by processing the body. Through stripping the meat from bone, the *Agent* ends up leaving a cut mark on the surface of the bone. The cut mark, in this example, thus becomes the *Trace*.

2.2. Taphonomy in the 21st Century

By the turn of the century, Taphonomy had undergone an enormous change, starting off as a relatively minor discipline in the grand scheme of archaeology, to becoming one of the protagonists in a heated scientific debate; the *Hunter-Scavenger* debate. The introduction of key concepts that included hominin populations as agents in taphonomic and cultural processes demanded a higher degree of scientific objectivity and empiricism in the classification of sites.

While the *hunter-scavenger* debate had a particularly notable importance towards the end of the 20th century, especially in anglo-saxon literature, casting a shadow of doubt over the legitimacy of some problematic interpretations of archaeological sites (Blumenschine, 1988; Marean et al., 1992; Blumenschine and Marean, 1993; Selvaggio, 1994a, b; Blumenschine, 1995; Capaldo, 1995, 1997), this debate was shook by a strong paradigm shift at the beginning of the 2000's with some key papers (Domínguez-Rodrigo, 1997; Domínguez-Rodrigo and Piqueras, 2003; Domínguez-Rodrigo and Barba, 2006; Domínguez-Rodrigo and Barba, 2007; Domínguez-Rodrigo et al., 2007). The importance of defining anthropic activity in a site became more than just identifying cut

marks, yet focused greatly on the identification of interactions between hominin populations and their surroundings.

Within this debate, the importance of studies regarding *Bone Surface Modifications* (BSMs) became increasingly apparent, supported by skeletal profiles, breakage patterns, and the systematic deconstruction of palimpsests in archaeological sites.

The order of meat consumption between carnivores and humans, for example, gave fruit to the interpretation of multiple other concepts, even reaching debates regarding the initial arrival of hominin populations in Europe (Turner, 1992; Arribas and Palmqvist, 1999). While reflections on carnivore behaviour can be dated back to the 1800's (Dawkins, 1863), early 1900's (Martin, 1906, 1907, 1909) and mid-20th century (Schaller and Lowther, 1969; Sutcliffe, 1970; Houtson, 1979; Brain, 1981) and considerable developments can be found in the 1990's (Cruz-Uribe, 1991; Capaldo and Blumenschine, 1994; Selvaggio, 1994a, b, 1998; Blumenschine et al., 1996; Ruiter and Berger, 2000), key research concerning the taphonomy of these agents, however, saw the 21st century revolutionise the study of carnivore accumulations, providing more in depth views into their behaviour and the traces they produced (Pickering, 2002; Pokines and Kerbis-Peterhans, 2007; Lansing et al., 2009; Bunn and Pickering, 2010; Pickering et al., 2011; Domínguez-Rodrigo et al., 2012a; Gidna et al., 2013; Yravedra et al., 2013, 2017c; Arriaza et al., 2016; Aramendi et al., 2017a). This provided more evidence that could be used against the defenders of the classic "*hominin scavenger*" models and began to truly unravel the trophic pressure in Eastern African Lower Pleistocene sites.

A means of studying carnivores is through the marks they leave on bones after the consumption of meat. Carnivores leave four main types of BSMs (Fig. 2.2.); tooth pits, scores, punctures and furrowing (Haynes, 1980; Binford, 1981; Brain, 1981; Haynes, 1983; Blumenschine, 1995). Tooth *pits* are product of the animal biting down, leaving a *circular mark* with a spherical base. If the carnivore were to then tear the meat away, and by chance draw the tooth across the surface of the bone, this is likely to produce a longer mark known as a *score*, with an equally rounded base. If the pit were to break the surface of the bone and penetrate the medullar cavity, this third main tooth mark would be called a *puncture*. The additional process of frequently chewing, and consequently destroying large areas of bone so as to access the medular cavity would be known as *furrowing*. Many carnivores leave specific patterns of *furrowing*, which are studied under the term *taphonotype/taphotype* (Wauthoz et al., 2003; Domínguez-Rodrigo and Pickering, 2010).

As for anthropic intervention in bone accumulations (Fig. 2.3.), the clearest direct evidence present can be seen in microscopic incisions on bone cortical surfaces. These are known as *cut marks*. A cut mark can be defined as a groove or *linear mark* that penetrates the surface of the bone through the displacement and movement of organic tissue (Walker and Long, 1977; Lyman, 1987; Potter, 2005). This is commonly produced by moving a sharpened edge against the bone's surface. While these marks are most commonly associated with lithic tools, they can also be produced with bamboo (Spennemann, 1990; West and Louys, 2007; Bonney, 2014), shell (Choi and Driwantoro, 2007; Weston et al., 2015), and metal tools (Greenfield, 1999, 2008; Bartosiewicz, 2009).

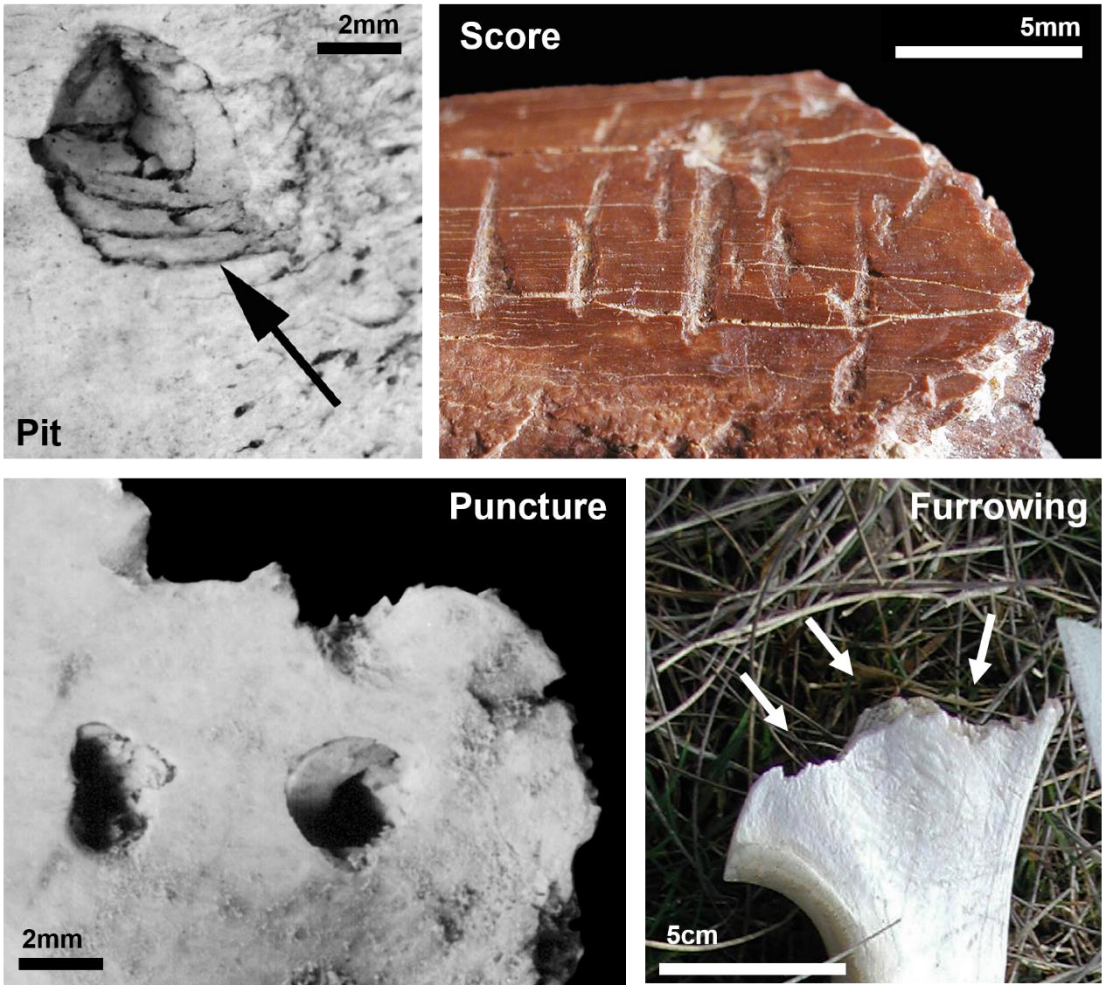


Figure 2.2 - Figures presenting the 4 most common types of carnivore damage to bone; pits, scores, punctures and furrowing. Photos from Fernández-Jalvo and Andrews (2016)'s Atlas of Taphonomic Alterations

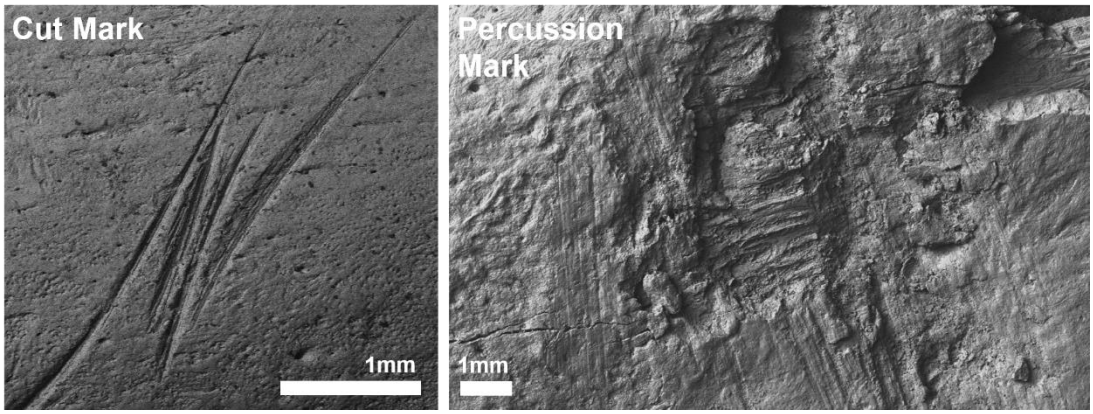


Figure 2.3 – Figures presenting most typical types of anthropic bone surface modifications; cut and percussion marks. Photos from Fernández-Jalvo and Andrews (2016)'s Atlas of Taphonomic Alterations

Other types of frequent anthropic BSMs are produced by breaking the bone to obtain the marrow (Noe-Nygaard, 1977; Myers et al., 1980; Blumenschine and Selvaggio, 1988; Villa and Mahieu, 1991). When this is carried out using a hammerstone, a *percussion mark* can be formed. Percussion marks are most frequently associated with blows to the cortical that have failed in breaking the bone. This impact mark leaves either a superficial circular pit, groove or series of microstriations (Pickering and Egeland, 2006).

The frequency, distribution, and nature of BSMs are then subject to detailed analyses that, in some cases, are able to discern between the different carnivore agents (Blumenschine, 1988; Capaldo, 1995; Domínguez-Rodrigo, 1999; Faith, 2007). Such data can then be considered highly important when understanding the order of consumption, and thus identifying which of the agents are more likely to be the hunter, and which of the agents is the scavenger.

Nevertheless, BSMs are not the only means of studying the taphonomic record of a faunal assemblage. In many cases, some sites unfortunately present bone cortical preservation rates that either partially or completely eliminate the taphonomic traces that would have been present prior to sedimentation (Pineda et al., 2014, 2019; Pineda and Saladié, 2018). If taphonomic research into the different agents consuming these carcasses were to solely rely on BSMs, these unfortunate cases would provide next to no data for the interpretation of the site. Fortunately, other factors can be considered such as the nature of bone fracture patterns.

When either the hominin or the carnivore achieves fracturing the bone to access the marrow, the mechanical nature of this fracture variates when considering the physical properties of the force breaking the bone (Green and Swiontkowski, 2000). When humans bring a hammerstone in contact with the bone, the physical impact is considered a dynamic process (i.e. dynamic loading). The pressure induced by a carnivore, however, when the teeth forces downwards on the bone until it breaks presents a completely different physical distribution of force. This is known as a static process (i.e. static loading). Both completely different physical processes alongside the natural morphology of the bone and cortical elasticity produce different patterns in the process of breaking the bone (Green and Swiontkowski, 2000). One major variable that summarises this is seen in the fracture plain angle (Pickering and Egeland, 2006; Blasco et al., 2014; Galán et al., 2009; Moclán and Domínguez-Rodrigo, 2018), and when a notch is presented, the dimensions of said notch (Capaldo and Blumenschine, 1994).

A combination of these different variables can be considered fundamental when deciphering the different agents and their role in the formation of an osteological assemblage. While, in many cases, their identification, description and definition had been originally published in the 1980's and 90's, most of these efforts into understanding the taphonomic register relied greatly on qualitative studies and simple observational techniques (Blumenschine et al., 1996). The 21st century, fuelled by a multitude of different scientific debates, brought more objective means of studying taphonomic processes with the use of more sophisticated analytical equipment, and the eventual incorporation of statistics in taphonomic studies.

2.3. Eastern African Taphonomy

Africa is the second largest continent in the world, and has the greatest diversity and abundance of mammal species (Kingdon et al., 2013; Kingdon, 2015). Eastern Africa alone, and the countries that lie along the Rift Valley (namely including Ethiopia, Kenya and Tanzania), contain up to 5 different biomes with vegetation ranging from the Acacia Savannas and the dry bushlands of the Serengeti to forest mosaics that eventually give way to the great expansion of the rainforests in the Congo basin. Product of such a diverse range of ecosystems presents over 127 different species of extant land mammals (Table 2.1.) over 12 taxonomic orders (Yalden et al., 1986; Wilson and Reeder, 2005; Kingdon et al., 2013; Foley et al., 2014). An additional 32 fossil species of bovidae are known alongside 77 non-bovidae species represented in the Olduvai Gorge alone (Leakey, 1965; A.W. Gentry and A. Gentry, 1978a, b).

Understanding the complexity of such a diverse ecosystem has proved difficult, with multiple stages of research development dating back to the era of exploration by European colonialists. Nevertheless, extrapolating this understanding of the African continent to the prehistoric record is still considered a topic of great complexity. Additionally, trying to understand the role of hominin populations and their evolution in this context is a highly problematic field of research (Bishop and Clark, 1967; Bunn, 1982a; Eldredge, 1989; Klein, 1989; Potts, 1996; Domínguez-Rodrigo, 1997; Diez-Martín, 2005; Domínguez-Rodrigo et al., 2007).

Order	**Family**	**Nº Species**
Tubulidentata	*Orycteropodidae*	1
Macroscelidea	*Macroscelididae*	6
Hyracoidea	*Procaviidae*	5
Proboscidea	*Elephantidae*	1
Primates	*Hominidae*	2
	Cercopithecidae	14
	Galagidae	12
Rodentia	*Pedetidae*	1
	Hystricidae	3
Lagomorpha	*Leporidae*	3
Erinaceomorpha	*Erinaceidae*	1
Pholidota	*Manidae*	3
Carnivora	*Canidae*	5
	Mustelidae	5
	Nandiniidae	1
	Felidae	6
	Viverridae	6
	Hyenidae	3
	Herpestidae	10
Perissodactyla	*Rhinocerontidae*	1
	Equidae	1
Artiodactyla	*Suidae*	2
	Hippopotamidae	1
	Giraffidae	1
	Bovidae	33

Table 2.1. – Table Presenting the Number of Species per Order and Family of Modern Wild Land Mammals in Tanzania. Data from: Wilson and Reeder, 2005; Kingdon et al., 2013; Foley et al., 2014.

NOTE: *Hominidae* includes the Genus *Homo*

Size Group	**Weight (kg)**
I	< 50
II	50 - 120
IIIa	120 - 200
IIIb	200 - 450
IV	450 - 1000
V	1000 - 4500
VI	4500 <

Table 2.2. - African Size Groups Adapted from Bunn (1987) and Bunn and Pickering (2010).

Our current understanding of early *Homo* origins and behavioural development regard the East of Africa as the origin of our evolutionary divergence, ca. 2 Ma (Wood and Richmond, 2002). This particular concept quickly received the title of the *East Side Story* (Coppens, 1994). Deciphering the context in which these populations existed requires a great deal of information obtained through multiple different approaches. Taphonomy, of which, can be considered a crucial component in East African evolutionary studies.

As can be imagined when considering such taxonomic and ecological diversity, East African taphonomy requires a particular methodological perspective that may vary from other sites in European contexts.

From one point of view, the great number of African Bovidae species can be considered highly important. The family Bovidae are the most diverse ungulate family to date, with over 270 extant species worldwide (Castelló, 2016). Africa, of which, is home to 86 (Kingdon et al., 2013); 76 of these divided into Antilopinae and 10 in Bovinae. When considering the country of Tanzania alone, a total of 33 extant species are currently listed, with a wide range of different shapes and sizes (Fig. 2.4.). These include everything from the Dik-dik (*Madoqua kirkii*) with an average weight of 5 kg (Foley et al., 2013; Kingdon et al., 2013), to the male African Buffallo (*Syncerus caffer*) reaching up to 850 kg (Kingdon et al., 2013).

Considering such a wide range of anatomical body weights, and such minute anatomical variations in many of the bovid species, many analysts face the difficulty of taxonomically classifying individuals when presented with highly fragmented accumulations. While in the case of bovids, many resort to a "lumper" classification approach according to taxonomic tribe, in cases where the species of the individual is unidentifiable, Brain (1974) proposed a means of efficiently dividing specimens according to size. Since this initial methodological approach, these weight classes have been revised numerous times and updated into 7 different size groups (Klein, 1976; Bunn et al., 1980; Brain, 1981; Bunn, 1987; Bunn and Pickering, 2010) (Table 2.2., Fig. 2.4. & 2.5.).

Group 1 consists of microfauna, rodents and many of the small forest antelopes (Cephalophini), dwarf antelopes (Neotrgini), Dik-diks (Madoqua) and most gazelle (Antilopini). Group 2 then includes the wild goats of Northern Africa, some smaller suids, as well as the remaining Neotragini, Cephalophini and Antilopini bovid species.

Group 3 is divided into 2 subgroups (Brain, 1981). This division was an original adaptation of Brain's original size group classifications from 1969 and 1974. This division proved successful when applied to the site of Klasies River Mouth (Klein, 1976), identifying differences in groups that had previously been lumped together directly as Group 3. This technique also adapts for differences some bovid species present through sexual dimorphism. Group 3a thus includes some of the larger suid species, both extant and fossil, hipparion as well as the smaller Tragelaphini species in the Genus *Tragelaphus*. In this group, other bovids such as the wilderbeast (Alcelaphini) and oryx (Hippotragini) are also included. Size Group 3b include zebra the rest of the Hippotragini tribe and the remainder of the slightly larger *Tragelaphus.* Among the *Tragelaphus*, the Greater Kudu and southern Waterbuck are included in the species that present great sexual dimorphism. In both cases the females are considered Group 3a while the males are included in Group 3b.

Group 4 is almost solely occupied by the African Buffalo and the Genus *Taurotragus*; last two members of the Tragelaphini tribe.

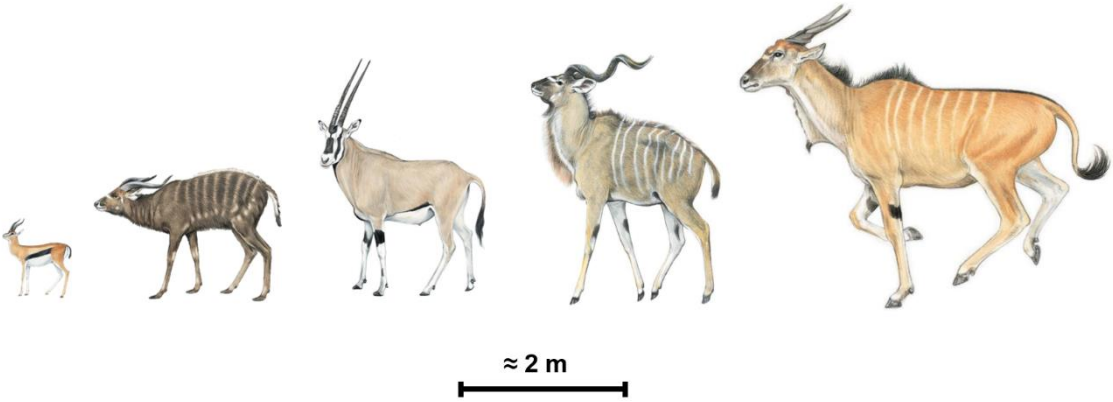


Figure 2.4. – Examples of African Sizes 1, 2, 3a, 3b and 4 (respectively) for African Bovids. All drawings are by Kingdon (2015) and are approximately scaled to show variation in size. Each species in order of appearance is Thomson's Gazelle (*Gazella rufifrons*), Sitatunga (*Tragelaphus spekei*), Beia Oryx (*Oryx beisa*), Greater Kudu (*Tragelaphus strepsiceros*) and Eland (*Taurotragus oryx*). All are frequent extant species in Tanzania and are also present in the fossil record.

Figure 2.5. – Some of the most representative animals of each size group (excluding antilopinae). In order of appearance: Porpucine (*Hystrix cristata*), Warthog (*Phacochoerus africanus*), Zebra (*Equus quagga*), Buffalo (*Syncerus caffer*), Hippopotamus (*Hippopotamus amphibius*), Elephant (*Loxodonta africana*). All drawings from Kingdon (2015). Drawings are NOT to scale. All are frequent extant species in Tanzania and are also present in the fossil record.

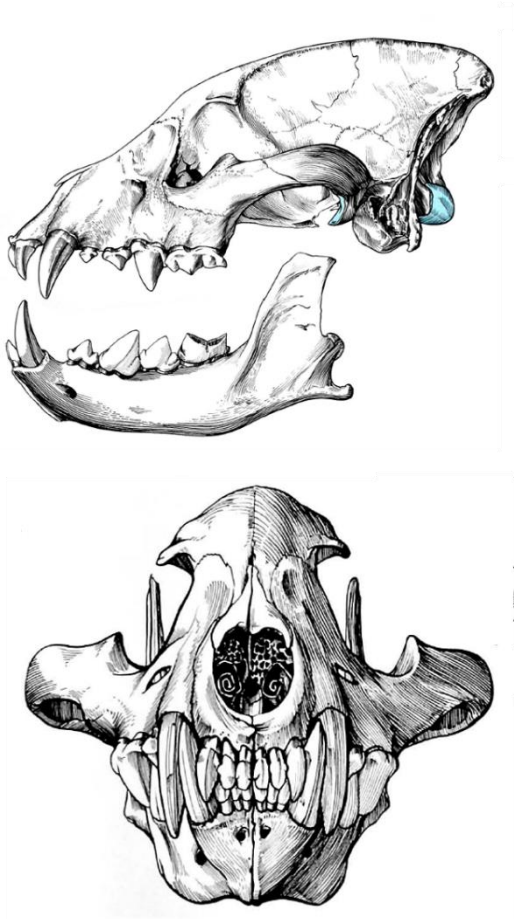


Figure 2.6. - Drawing of the Cranial Features of a Hyena's Skull. Source (Pales and Lambert, 1971

Size 5 consists in the largest mammals of the Animal Kingdom, and thus considered Megafauna; weighing over a tonne. These include the giraffe, rhinoceros and hippopotamus. Among these species, the fossil record presents a great deal of larger species that also belong in this group, including the *Sivatherium* of the Girrafidae family (Harris, 1976), and the much larger Pelorovis, of the Bovini tribe (Gentry, 1967).

Size 6 is exclusive to Proboscidea, considered too big an animal to be lumped with any other African mammal.

Africa is also home to a total of 83 Carnivore species that can be divided into Caniforms (Families = 4, n = 25) and Feliforms (Families = 5, n = 58). The four most important carnivores, especially in association with African archaeology, are the Hyena, Lion, Leopard and Wild African Dog.

Hyenas have shared the same landscape with hominin populations throughout the majority of the Pleistocene in Africa and Eurasia (Brain, 1981: 56). Hyenas, much like humans, are also keen accumulators of bones with the capacity to heavily alter a faunal assemblage through their consumption of animal carcasses. Unlike other animals, such as leopards (Brain, 1981; Ruiter and Berger, 2000; Martín, 2008; Domínguez-Rodrigo and Pickering, 2010; Pickering et al., 2011) and wolves (Binford, 1981; Haynes, 1983; Esteban-Nadal et al., 2010; Yravedra et al., 2011; Sala et al., 2013), hyenas normally produce a high degree of fragmentation alongside the destruction of bone surfaces (Kruuk, 1972; Cruz-Uribe, 1991; Selvaggio, 1998; Pickering, 2002; Domínguez-Rodrigo et al., 2012a), making these animals one of the most destructive taphonomic agents in the paleontological/archaeological register.

One of the main reasons behind the hyena's destructive tendencies is found in their durophagus nature. This is especially evident in at least 3 of the 4 extant hyena species; *Parahyena brunnea*, *Hyaena hyaena* and *Crocuta crocuta*. From a paleontological perspective, these animals have undergone millions of years of evolution in which they have developed key anatomic features that contribute to their bone-cracking capacity (Koepfli et al., 2006). These features consist primarily in their cranial morphology as well as their dentition (Fig. 2.6.). Osteologically, these animals present a robust zygomatic arch, powerful saggital crest, shorter snout and a stepped frontal profile, built to relieve mechanical stress when crunching bone (Pales and Lambert, 1971; Bonifay, 1996; Ferretti, 2007). Additionally, their dental configuration consists in a combination of sharp elongated upper dentition and robust lower dentition (Hillson, 2005).

Furthermore, a distinct development of the upper premolar metaconid (except in *Crocuta*), as well as more developed upper incisors, are ideal for bone cracking (Ferretti, 2007). Considering enamel structure, the distinct zig-zag nature of hyaenid enamel configuration ensures a reduced risk of shattering the tooth upon occlusion (Ferretti, 2007). All of these features combine to create the hyena's perfect bone-cracking masticatory system.

The bones left by hyena has proven to be a difficult animal to study. While certain patterns have been observed through studies of taphotypes/taphonotypes (Domínguez-Rodrigo and Pickering, 2010; Domínguez-Rodrigo et al., 2012a, 2015b), and other markers such as tooth pits and scores, hyenas produce such a wide range of alterations that sometimes deciphering other taphonomic agents is much more difficult (Hughes, 1954; Cruz-Uribe, 1991, Pickering, 2002, Pokines and Peterhans, 2007; Prendergast and Domínguez-Rodrigo, 2008). While other analysts have tried to use tooth pit and score dimensions (Selvaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003; Delaney-Rivera et al., 2009; Andrés et al., 2012), as a method in discerning carnivore agency, statistical overlaps create issues in interpretation and can lead to great confusion. Some qualitative observations of tooth marks, especially in the case of crocodiles (Njau, 2006; Njau and Blumenschine, 2006; Baquedano et al., 2011; Westaway et al., 2011; Njau and Gilbert, 2016; Sahle et al., 2017), are powerful tools, however, this method is still highly susceptible to other variables that can create confusion and cause misdiagnosis (Sahle et al., 2017).

In general, hyaenids can taphonomically be characterised by an intense frequency of tooth marks with no preferential selection of size nor taxa in their prey. Hyaenids are great accumulators of bones, yet are known for transporting mostly appendicular elements. Concerning taphotypes, these animals often leave cylinders.

From this perspective, zooarchaeological quantitative analysis has also tried to confront these issues through the degree of fragmentation and skeletal part representation. Some authors, taking into account bone density studies (Kreutzer, 1992; Lam et al., 1998; 1999; 2003), have made an attempt at looking into how hyena intervention may affect Number of Identifiable Specimens (NISP), Minimum Number of Elements (MNE) and Minimum Number of Individuals (MNI) counts (Marean and Frey, 1997; Pickering et al., 2003; Faith et al., 2007). Despite these reflections, the differentiation between accumulative agents is still fairly problematic.

One of the two African Felids, Leopards (*Panthera pardus*) are peculiar accumulative agents, presenting a particular behavioural habit of protecting and storing their prey up in trees (Brain, 1981). To a lesser degree, leopards have also been observed to hide their prey in caves (Peterhans and Singer, 2006). Leopards, contrary to hyenas, are known for transporting complete carcasses (Peterhans and Singer, 2006; Pokines and Peterhans, 2007), and leave fewer BSMs than hyenas (Ruiter and Berger, 1999). They have also been observed to produce catastrophic age patterns, preferentially hunting adults individuals over juveniles (Bunn and Pickering, 2010).

Lions (*Panthera leo*), on the other hand, are quite different. Lions are the largest African carnivore, acting as a keystone in the ecosystem[1]. These felids,

[1] They play a larger role in the ecosystem than can be expected when considering the abundance of said species.

contrary to Leopards, tend to consumer their prey at the kill site, yet have recently been found to be accumulative agents as well, as observed in the Olduvai site of OCS (Arriaza et al., 2016). Much like leopards, these felids leave few BSMs and the alterations they produce to the bone overall is much less to that of a hyena (Fig. 2.7.) (Domínguez et al., 2012a, 2015; Gidna et al., 2015). Because of their size, however, and their social hunting tendencies, lions are able to hunt much larger animals and often go for individuals over African size Group III.

The African wild dog (*Lycaon pictus*) is a gregarious animal that preys on a wide range of different animals. Little work, however, has been done on this particular animal's activities and possible role in the taphonomic register. The African wild dog is known to be an agent that moderately alternates bones, with less consumption of the actual bone and more consumption of meat (Yravedra et al., 2013).

Contextualising this information with the other common carnivores of the African savannah, the Hyena can be considered the most active and damaging taphonomic agent to osteological assemblages, followed by canids and. finally felids. Nevertheless, taphonomy is in a constant state of development, and more research is needed to unravel the true nature of many archaeological sites. One key issue frequently being presented in all cases of archaeological research is that of *Equifinality*. Equifinality is a fundamental concept that continues to trouble archaeologists to date. It's presence in the taphonomic research leads many to withdraw erroneous conclusions and problematic interpretations, highlighting the need to focus our efforts on overcoming this issue; presenting one of the key components of this paper.

Figure 2.7. - Intermediate Appendicular Elements eaten by Felids (above) and Hyenids (Bellow). Source: Domínguez-Rodrigo et al., 2015

2.4. Equifinality

Equifinality can be defined as the likelihood of misclassifying a taphonomic process or defining two completely different traces as one and the same. If two systems are indistinguishable through taphonomic analysis, then we are unable to withdraw hypotheses from them and are likely to misinterpret the nature of a trace or process.

The 21st century in taphonomy can be seen to have provided new perspectives in which to confront taphonomic equifinality, especially with regard to carnivore activity. Considering the trophic pressure of the African Savanah and the competition for resources between hominins and carnivores, many taphonomic efforts have relied heavily on new statistical and analytical approaches to separate

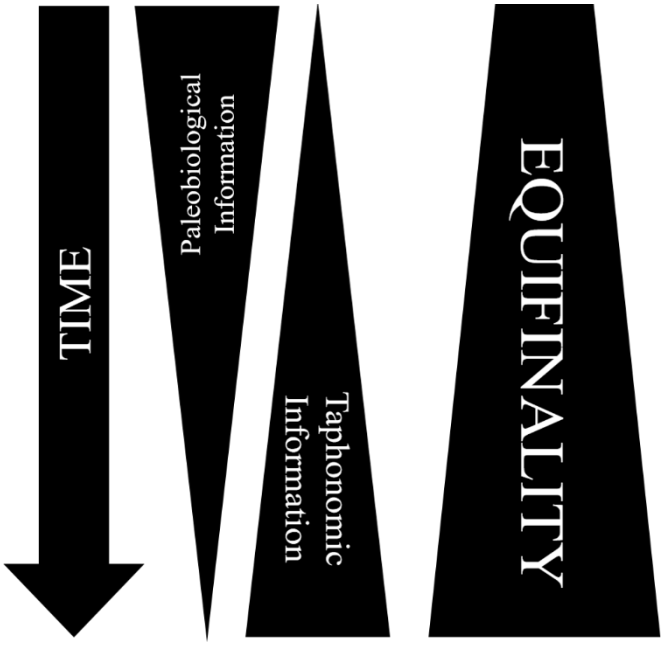


Figure 2.8. - Diagram Representing the Increase of Taphonomic Equifinality with the Passing of Time

different accumulative agents. However, when a site presents multiple agents acting on the same bones, at times it becomes very difficult to separate the activity of one agent from another. This is especially evident in moments where hominin-carnivore interactions surround hunting and scavenging patterns.

To further complicate this topic, many post-depositional processes are able to produce taphonomic traces as well, and in many cases, erase the taphonomic traces that were already present before them. If we were to reconsider the previous mentioned work of Fernando-López (1991: Fig. 2.1.), it is logical to assume that with the increase of the taphonomic register and the decrease in the faunal record product of time progressing, the likelihood of marks overlapping increases greatly. To develop this matter, as certain more destructive marks overlay previous traces, it is equally as likely that the physical nature of these marks change. In some cases to the point of elimination. This constant development of taphonomic overlay is perfectly defined by equifinality, and can be produced at any point of the taphonomic chain. If we are then to consider how a great quantity of data is already subject to equifinality prior to sedimentation (such as the tooth marks produced by different carnivore agents), the visualisation of the entire process can be described as Fig. 2.8..

After death, a number of processes occur to the carcass of any living organism. If these remains last long enough, the process of sedimentation will eventually begin. Even though the bone is covered and thus protected from further modification by external agents, the process of decay and any other diagenetic biochemical interactions between the bone and its surroundings are capable of producing marks BSMs.

One possible agent to be considered is the presence of vegetation. Plant roots naturally grow searching for nutrients they need to survive. Associated with this organic material are multiple fungi and bacteria that react chemically with the surface of the bone, producing a series of different modifications to the bone's surface (Fernández-Jalvo and Andrews, 2016). These may include staining or corrosion, known as *dendritic etching* (Andrews and Cook, 1985). A famous case of equifinality regarding *dendritic etchings* can be found in their misinterpretation as tooth marks by Blumenschine (1986, 1988). Such confusion formed part of the heated hunter scavenger debate, and was resolved through the use of high resolution analytical techniques

(Domínguez-Rodrigo et al., 2007: 39-100), including more advanced microscopic analysis than relying solely on the use of a hand held lens.

The issue with simple characterisation of taphonomic traces in this manner, however, occurs when marks show very similar physical and structural properties. For example, a percussion pit and a tooth pit in many cases have almost the exact same structural appearance. They are both circular in nature, they may or may not be associated to microstriations and in many cases their internal morphology can appear identical (Pickering and Egeland, 2006; Galán et al., 2009). If from a qualitative perspective these marks can be confused, then new more rigid criteria are required for their differentiation.

One of the greatest taphonomic agents to create equifinality with anthropic activity, however, is the sediment itself. *Trampling* marks are a phenomenon produced by sedimentary abrasion across the surface of the bone. These marks are generally attributed to pressure being exerted across the osteological surface, thus forcing sedimentary particles to scratch and make an incision across the cortical. In appearance, these marks can be described as superficial, irregular and _/ shaped incisions (Brain, 1967; Behrensmeyer, 1978, 1986; Andrews and Cook, 1985; Fisher, 1995), however, depending on the sedimentary context and time exposed to the effects of trampling, their size may vary in shape and depth.

After the initial discovery of trampling marks, issues with their definition arose especially throughout the 1980's (Fiorillo, 1984; Oliver, 1984; Andrews and Cook, 1985; Behrensmeyer et al., 1986). One of the greatest difficulties related with these observational studies, however, can be considered in the experimentation involved. Without truly understanding the phenomena responsible for trampling marks, authors were unable to truly associate marks with the process involved in their formation (Domínguez-Rodrigo et al., 2009a). While Olsen and Shipman (1988) published trampling marks to be absent of microstriations, Behrensmeyer et al. (1986) published the opposite. Carefully regarding the experimental protocol presented by these authors, however, reveals how the nature of these trampling marks may be product of other features. These variables may include the time exposed to abrasive activities or the weight the bones were subjected to during said activities.

With more advanced means of assessing BSMs, statistical approaches and developed experimental protocol began to reveal new means of differentiating between these types of marks (Domínguez-Rodrigo et al., 2009a; Juana et al., 2010). Nevertheless, these approaches still relied heavily on qualitative observations, subject to analytical subjectivity. Autocritical revisions of these papers revealed how the nature of the variables used, as well as the analyst's experience (Domínguez-Rodrigo et al., 2011b), are likely to produce large margins of error in BSM analysis (Domínguez-Rodrigo et al., 2017b). To some degree, other authors have even come to observe similar variables produced by other agents (Sahle et al., 2017), increasing the presence of equifinality in linear mark classification. Similar issues are increasingly evident in the archaeological register, where other taphonomic processes such as chemical alteration and fluvial abrasion are able to alternate the nature of BSMs, erasing features used for their classification, or even eliminating them completely (Pineda et al., 2014, 2019; Pineda and Saladié, 2018).

The presence of equifinality is an ever increasing issue in the study of archaeological accumulations, producing increasing debate regarding the interpretation of sites. To name a few, heated debates currently argue the reliability of evidence claiming the butchering practices of *Australopithecine* populations in Dikika (McPherron et al., 2010, 2011). While some argue the presence of cut marked bones (Thompson et al., 2015), other authors call for a more rigid experimental approach when processing such controversial remains (Domínguez-Rodrigo et al., 2010; Domínguez-Rodrigo et al., 2011b; Domínguez-Rodrigo and

Alcalá, 2016). Likewise, 130 Ka butchered bones from America (Holen et al., 2017) have also been called to question (Domínguez-Rodrigo, 2018). The majority of these studies rely heavily on highly subjective traditional means of processing BSMs (Domínguez-Rodrigo et al., 2017b), with questionable statistical means of providing a significant sample that supports these claims (Domínguez-Rodrigo, 2018). What becomes increasingly evident is how our current means of identifying and classifying BSMs remains highly subjective. Especially in sites where the association of faunal and lithic remains are unclear, and the implications of these finds can be considered highly controversial.

In order to evolve, taphonomy needs to heavily rely on some key components, including new statistical approaches and high definition means of processing data. Nevertheless, using carefully planned experimental protocol, supported by new advances in microscopy and advanced statistics, we hope to confront the taphonomic register from a different perspective.

2.5. Epistemology and Archaeological Science

It can be considered that one of the greatest issues taphonomy has had to face is the need for empirically scientific protocol in its practice. Throughout the 20th century, theoretical debates ensued within archaeology to argue its classification or not as a science. For many, this can be considered to be an issue between processual and post-processual schoolings (Renfrew and Bahn, 2016; Johnson, 2009). While some have argued the need for scientific methods in our research (Renfrew, 1973a, b; Binford, 1981; Shanks and Tilley, 1987a, b; Renfrew 1989), others have considered that objectively classifying human populations in such a manner is not true to our basic nature, nor the correct means to approaching the past (Hodder, 2008). For the field of taphonomy however, a means of systematically processing data becomes a necessity for any type of interpretation (Binford, 1981), requiring alongside it a series of theoretical concepts that define this field of research.

From one perspective, it is important to consider that archaeology is not a singular field, yet a combination of multiple disciplines with one common objective. Since the beginning of its conception, archaeology has borrowed methodology from other fields of research, employing both theoretical and methodological techniques from fields of research including geology, biology, chemistry and even mathematics. To some degree, the earlier practitioners of archaeological research weren't even considered archaeologists themselves, considering how the initial discoveries of the Acheulean technocomplex were achieved by apothacaries and doctors in medicine (Hearne, 1770).

To this extent, it can be seen how archaeology is a discipline constructed from many other disciplines. Taphonomy in itself is a product of paleontological schoolings, however archaeology has adapted this field of research to benefit its own form of knowledge. A large part of medieval and ancient history, for example, is product of literary texts and to some extent can even be told orally (Holtorf, 1998; González-Álvarez, 2011), however when considering the first populations of the Eastern African rift valley, a clear lack of direct sources makes approaching prehistory in this manner impossible. There is no direct means of knowing the nature of *Homo habilis* populations 2 Ma, therefore prehistoric archaeology has relied heavily on other scientific disciplines to develop a specific tool kit for the research of these populations.

A fundamental issue with this concept, however, is a distinct lack in updated knowledge that contributes to our research. Stapert and Street (1997), when referring to spatial archaeological statistics, stated that "multivariate procedures have hidden mathematical assumptions which are not well adapted to archaeological data". What these authors fail to realise, however, is how science is a non-linear construction that is constantly evolving to accommodate the problems we have to face on a daily basis. While authors within archaeology have borrowed methods from fields of biology, for example, and have repeatedly used these methods. Stapert and Street (1997) additionally fail to realise that the field of biology is constantly adapting, and the same methods we repeatedly employ in investigation, to many biologists, have now become obsolete. Not all scientific theories are necessarily laws, and most are prone to change. This requires any scientific field to frequently publish their advances, update the current knowledge on their given topic and provide new means of acquiring said knowledge. In many cases this requires in depth collaborative projects, whereby multiple specialists work together to provide new means of processing data.

To respond to Stapert and Street (1977), therefore, is to argue that the issue is not with the "mathematical assumptions", yet with the analyst employing these techniques in their study. Any archaeologist using a tool from another field, should be aware and reflect not only on the benefits of this technique, but also on its limitations. In many cases, a statistical test may have been adopted to respond to a specific question (Ripley, 1977, 1979, 1981), however with the passing of time, mathematicians have reflected on the fundamental concepts of this test, thus updating it for other applications (Baddeley et al., 2000; Dixon, 2002). As in any field, a scientist should always be aware of the advances happening around them, ensuring to remain flexible to the non-linear nature of every discipline (Kuhn, 1962; Feyerabend, 1993).

This is especially evident when considering basic functional concepts such as the creation of a cut mark. In reality it is impossible to understand the mechanisms of the mark if we are to ignore the tool producing it. Many authors have argued that raw materials and tool type are conditioning factors in mark features and morphology (Domínguez-Rodrigo et al., 2009a; Maté-González et al., 2015; Moclán et al., 2018, e.g.). Therefore a taphonomist should not isolate the final product without considering the process behind its creation. To this degree, it may be important to consider the work of geologists and lithic tool experts when confronting these contexts.

Acquiring knowledge and the consequential sharing of this information is fundamental in scientific research (Mazloumian et al., 2012), and frequent malpractice within multiple fields can be seen in the arbitrary reuse of old techniques without considering what other fields of research may have to offer.

Epistemologically speaking, archaeology fulfils many of the criteria to consider itself a science, especially considering the employed methodology in multiple cases. However, the primary issue lies in the development of this field. As has been discussed, taphonomy for many years has faced numerous problems with equifinality, especially considering the lack of agreeable definitions for the formation and features of naturally produced trampling marks. The main issue to be considered, however, is the protocol employed when confronting these questions and the subjectivity behind the selected criteria used in evaluation. Experimentation, for example, is a fundamental component of any science, however the means of approaching experimentation can be considered poor throughout the majority of archaeological research over the past couple of decades. Archaeology, as any science should, needs

to evolve. The primary means of evolving, however, lies in the definition and correct use of scientific protocol, such as that of experimentation.

2.6. Experimental Archaeology

One of the most important sources of information, especially in taphonomy, lies in experimentation. Epistemologically, scientific investigation relies on three basic concepts; induction, abduction and deduction (Eco and Sebeok, 1989). These three concepts orientate themselves around our own experiences, our observational capabilities and the known facts at hand:

1. Deduction consists in the recreation of an experience in order to form a hypothesis explaining our observations.
2. Abduction is based on an observation of a phenomenon, resulting in the search for the conditions that essentially explain said observation.
3. Induction is used to search for sufficient observations that can directly define and support a hypothesis, relying heavily on our current knowledge.

On this basis, if we were to recreate the "experience" in which archaeological phenomenon are produced, our observations would essentially *deduce* the hypotheses; thus defining the main objectives in experimental archaeology. Upon our deductions we are able to generate a basis for which we can abduce and induce various other factors that appear within the archaeological record.

In order to properly explain the art of deduction within experimental archaeology, we have to understand that a great number of variables may be affecting our results. In order to successfully withdraw hypotheses from our experimentation we have to be able to eliminate all other possibilities in order to be sure that our hypothesis is correct (Doyle, 1890). This concept essentially defines falsifiability as a *Criterion of Demarcation* (Popper, 1937).

This concept archaeologically can be relatively summarised through the *Middle-Range Theory*. The concept of the Middle-Range Theory was first adopted by the American sociologist Robert Merton (1967, 1968), in his attempt to integrate theory and empirical research. Merton, inspired heavily by the work of Talcott Parsons (1968, 1975), argued that general statements could be verified through the understanding of certain aspects of social phenomena. Merton argued that trying to understand an organism as a whole was prone to subjective theorizing, whereas if each aspect of a phenomenon is considered separately, using measurable aspects of social realities, then a theory could be constructed rather than broadly assumed. In sociology, this concept relies on understanding human action in conjunction with the different motivational components that produced the final product (Parsons, 1975).

The application of the Middle-Range Theory to archaeology, however, wasn't properly developed until the work of Lewis Binford (1967; 1968; 1981), during the initial phases of Processual Archaeological theory (Johnson, 1999; Atici, 2006). Binford developed the concept of scientific analogy in archaeology while developing methods for archaeological deduction (Binford, 1967, 1981), highlighting the importance of analogy between elements in order to construct accurate archaeological theories. Binford's main argument lies in the [deduced] construction of multiple hypotheses, enhancing our understanding of observations. Through the construction of multiple theoretical frames of reference (Gifford-Gonzalez, 1991), we can begin constructing theories on hominin behavioural activities of the past (Fig. 2.9.). This becomes alarmingly apparent when observing realities of the modern world in

conjunction with our observations of the past (Atici, 2006).

While Binford based the majority of his Middle Range Theory research on ethnoarchaeological studies of modern Nunamiut populations (Binford, 1978, 1981, 1984a), the last couple of decades has seen a significant increase in experimental studies focusing specifically on the use of experimentation to understand various elements observed in the fossil record.

The use of modern day experimentation in taphonomy is a highly important factor in our understanding of how certain phenomena work, provided through our ability to directly observe the cause and effect of multiple trace marks and the agents that produce them. To provide an example; after observing with our own eyes how a carnivore consumes its prey, we can then study the tooth marks produced in a new light. This concept has proven very useful when considering the feeding habits of carnivores such as the crocodile (Njau, 2006; Njau and Blumenschine, 2006; Baquedano et al., 2011; Westaway et al., 2011; Njau and Gilbert, 2016; Sahle et al., 2017; Domínguez and Baquedano, 2018). The same can be said for lithic tool production (Schick and Toth, 1998; Dibble and Rezek, 2009; Geribàs et al., 2010; Moore and Perston, 2016); in attempting to produce flakes ourselves we can appreciate the difficulties and complexity of an activity that most would find incredibly simple: it isn't just a case of bashing rocks together (Turner, 2013). Experimentation is a powerful tool that we can utilise to produce new ways of studying certain activities. In constructing multiple observed hypotheses from experimentation, we are providing theoretical references to construct the Middle Range Theory about hominid behaviour in the past.

This same concept has been used to aid a significant proportion of the academic fields related to studying human cognition and its evolution.

Current understanding of hominin evolution and the necessary cognitive capacities required to knap indicate a strong need for social interaction, seen through recent studies into the advantages of social learning. Through experimentation (Lombao et al., 2017; Petö et al., 2018), observational studies into the neuroscience of learning in humans (Stout and Chaminade, 2007, 2012; Stout et al., 2015), and chimpanzees (Tomasello et al., 1987;

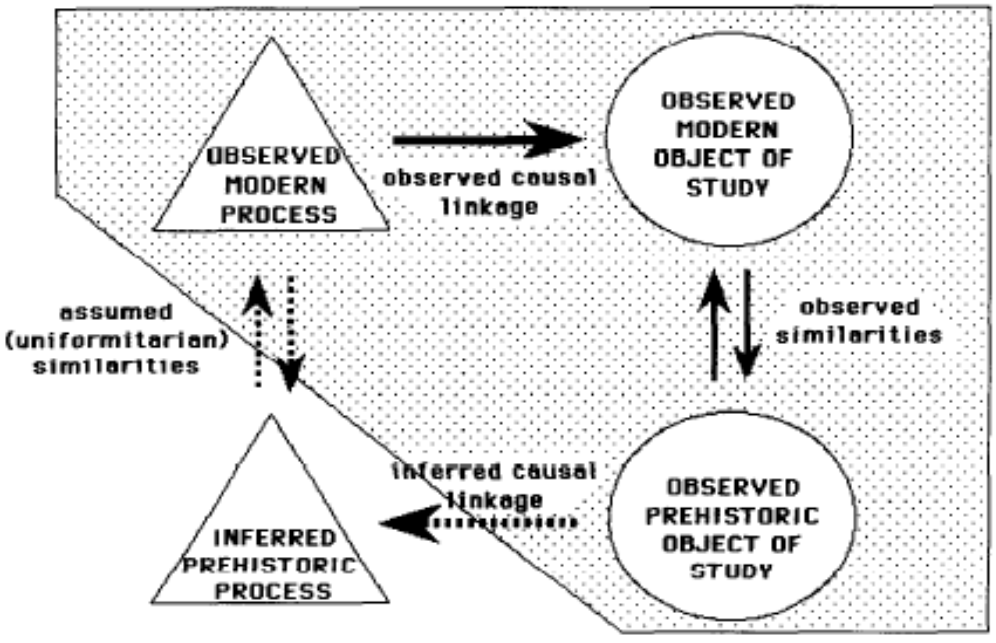


Figure 2.9. - Model proposing the logic of archaeological investigation through experimentation (Gifford-Gonzalez, 1991:222)

Matsumura, 2004; Melis et al., 2011), experimentation and actualistic studies have been valuable tools in the construction of many theories regarding cognitive evolution and the associated behavioural attributes of extant hominin species.

A significant problem with experimental archaeology, however, is that not all experimentation is valid. Many archaeologists (especially as students) tend to lean towards the attractive side of experimentation without taking into account the scientific implications of performing a valid experiment. This is most probably due to a lack of scientific formation in the humanities. As a result of this, many novice archaeologists are unfamiliar with the basic concepts behind scientific logic and the formation of valid theories.

In science, the concept of a "fair-test" is fundamental and basic (Popper, 1935, 1963), even for primary school students (Department of Education (U.K.), 2013). A fair-test implies that the elements we are observing are standardised and absent of bias. Within this concept we define a simple establishment of dependent and independent variables. A dependent variable is, as its name suggests, dependent on the other 'independent' variables. Upon changing the independent variables we are able to observe the cause and effect of each variable separately; successfully withdrawing a hypothesis on how each variable conditions the experiment. This is a very basic concept that many fail to understand, or even choose to ignore. To take this a step further it is necessary to talk about analogy and equifinality.

Analogy, as defined by Bunge (1981), is the comparison between multiple systems. These systems can consist of either qualitative or quantitative elements, depending on the investigation objectives at hand. If we focus on the second definition as proposed by Bunge, two systems have to have a composition that is substantially, environmentally or structurally analogous. The degree of analogy can be calculated through a number of logical conditions. However, in our case, what we need to draw attention to is the necessity to ensure that the dependant variables we are comparing are similar. The independent variable is the element which we should be controlling, considering how, if there are multiple variables affecting the results of our experiment, the value of our results thus becomes useless. What this calls for is the careful planning of any experiment to ensure our observational capabilities in order to see how each variable can affect the results we are trying to obtain (Popper, 1935).

Equifinality, on the other hand, is defined as an end state which may have been produced by two different potential means. If two systems are exactly the same or indistinguishable through statistical analysis, then we are unable to withdraw hypotheses from them (Bertalanffy, 1949, 1968).

The combination of these two concepts highlights the necessity for careful planning when preparing for an experiment. In order to construct an experiment with valid results, we have to be able to scientifically control the conditions under which the experiment takes place. In order to do this, we need to be able to reduce statistical noise by limiting out observations to one variable at a time. The overall results and observations produced by changes to said variable can thus explain the effects an element has on our conclusions. This planning, however, has to be as precise as possible in order to limit our parameters and observe the effects on our final outcome (Popper, 1935).

The work of Lyell (1830) developed the concept of uniformitarianism (Hutton, 1794; Playfair, 1802; Whewell, 1847); defining our natural assumption that the same scientific laws of nature that we apply today are applicable universally to everything else in time. This metaphysical and geological concept that "the present is the key to the past" ties well with archaeological theory in a very evident way. To provide the rather ridiculously simplistic, yet logical, example that gravity has always existed throughout our time on earth; we must also consider that the physical mechanics

and properties of stone tools have also remained consistent. The mechanical process of knapping a flint core is constant, we must also assume, in as such, that the physical action of passing a sharp blade across osteological surfaces will also produce the same results now as they would have done 400,000 years ago. Modus ponens, if planned correctly, our experimentation may provide a theoretical window into the past of our species, including a more personal experience rather than a simplistic theoretical understanding. In combination with this we therefore have to ensure analogy between the systems we are testing, the conditions we intend to produce, therefore, have to be as similar as possible to the conditions that were present during the production of the sample we are studying. Through this we generate controlled actualistic observations from which we can build, test and prove the hypothesis we wish to argue.

If we are truly able to understand the processes that generate a particular taphonomic trace, we can begin to decipher the mechanisms that interact with the fossil record and thus interpret these diagnostic traces that are left behind (Atici, 2006). Needless to say, this is quite an optimistic attitude towards science and is unfortunately never as easy as it sounds.

In reality, through statistical studies into the comparison of samples, an important scientific concept to consider is that of falsifiability. Falsifiability as a criterion of demarcation (Popper, 1935: 17-20) states that in order to create a "valid" statement we have to reject all possibilities that the statement cannot be true. Experimentation in as such should be a means of determining the dissimilarities of variables rather than a means of deciphering what they have in common. In taphonomy, the objective therefore is to highlight these differences, and thus reduce the subjective equifinality in our observations.

A pessimistically daunting reflection on this, however, defines a hypothesis as infinitely undeterminable. Through proposing new methodologies in experimentation, and with the constant advances produced through technological developments, we are always able to improve the reliability of our results (Popper, 1935: 94). Unfortunately, such is the rule of science.

Nevertheless, this same pessimism can be reversed through our determination, dedication and our understanding of *investigation*. We define investigation as the formal or systematic examination of research (Oxdord Dictionary, 2018). In trying to solve a scientific problem, for example, and consequently realising our incapability of providing a definite truth; we open ourselves up to further possibilities of discovering a means to an end.

Experimental archaeology is a fundamental tool within taphonomy, however, we need to establish the correct criteria in order to define what we are studying. Through Middle-Range theories we can define archaeological sites, analyse the remains and assign meaning to their uses (Gifford-Gonzalez, 1991). Through the organisation and planning of different experiments we are able to conduct valid research and generate the multiple frames of understanding that are necessary in deciphering these traces.

2.7. Microscopy

The origin of microscopy roots back to the "Age of Discovery", where magnified lenses were used by the likes of Galileo and Drebbel for the study of small objects (Hurerta, 2003). Their eventual development for studies in microbiology and science can be dated to the XVIIth century with the work of Antoine van Leeuwenhoek (Nick, 2015).

Microscopy in essence is the use of multiple lenses and their consequent reflection of light to magnify the object under observation. The physical properties behind the formation of the image requires

the optical system to be composed of different refracting surfaces, ensuring that the ray or beam of light passes through the surface of the lens.

Initial developments in microscopy aimed at using these magnifying optical systems to reveal details of objects that could not be viewed through simple observations and the naked eye. The resulting "microscopes", however, consisted quite simply in a singular lens that was able to provide a more detailed view of the object under study. The scientific paradigm shift that eventually revolutionised these optical systems was found in a combination of multiple convex lenses, thus producing higher magnifications (Abramowitz and Davidson, 1999).

With more developed technological and mechanical developments throughout the XIXth century, technicians and manufacturers began combining developed knowledge in physics to provide more complex optical systems, combining both the qualities of reflecting and refracting surfaces to produce higher quality images in more developed compound microscopes. Needless to say, come the XXth century a high number of different optical pieces of equipment were at the disposal of scientists with the ability of reaching incredible degrees of magnification.

Microscopy has been the focus of study for many generations, with new advances in this field aiming to provide more powerful and resolute imaging of micro structures and organisms through combined use of lenses and different energy waves. In general, microscopy can be divided into several subfields: Optical, Electron, Scanning Probe and X-Ray microscopy.

Optical Microscopy employs the portion of the electromagnetic spectrum that is visible to the naked eye in order to provide a magnified view of a sample. While traditional microscopes require direct observation, new technological advances have allowed for combined usage of digital processing of these images, providing real-time images of the sample under inspection.

With increasing advances in optical physics came the inclusion of techniques that employed different wavelengths of light, overcoming traditional optical techniques by providing new types of visual studies. Arguably the largest development in this field can be considered the development of Scanning Electron Microscopy (SEM) and Transmission Electron Microscopy (TEM) that have been able to reach resolutions of up to nanometres. Resolution is determined essentially by the wavelength of photons projected onto the sample, thus the use of electron waves as opposed to photon waves have been able to overcome the 0.2 µm resolution produced by standard optical microscopes (Goldstein et al., 2003). To a similar effect, other techniques have tried to use advances in X-ray beams that are able to provide higher resolution than either photon or electron waves.

So far, the aforementioned developments in the inspection of microscopic surfaces have involved non-invasive non-contact techniques, employing the use of lenses and light energy waves to inspect the object under study. Nevertheless, contact techniques such as Scanning Probe microscopy do exist, employing the use of a probe to track the microtopography of almost flat surfaces (Salapaka and Salapaka, 2008).

With such a wide number of different analytical systems, choosing the right microscope and analytical technique requires a means of evaluating the effectiveness and benefits of each piece of equipment. A major component of this consists in understanding the processes involved in image formation.

When observing the microscopic features of a given surface, the nature of particular surfaces may prove problematic when using standard visible light, especially if said surface is reflective in nature. Overcoming this issue in certain

circumstances can be obtained using polarised filters. Polarised filters provide a means of filtering the amount and *type* of light that passes through, and thus helps when searching for contrast or variation in particular surfaces (Maksymilian, 1989). Different polarisation techniques exist, yet the objective is to ensure that all light waves are vibrating along the same plane, thus enhancing our observational capacities of certain substances.

When considering the nature of a lens, we must also consider the degree to which it *refracts* a ray of light. The phenomenon of refraction can be described by a means of changing the direction of a ray of light, depending on the properties of the medium in which it penetrates. When the light wave passes from one medium to another, the speed of the wave is altered depending on the properties and refraction index of each medium as well as the angle of which the ray is introduced to this change. This can be described by *Snell's Law* (Makysmilian, 1988). In microscopy, the spherical shape of the lens as well as its refractive index conditions the means in which the light passes through. In an optical system, such as that of a microscope, the use of two or more lenses can enhance image formation and thus achieve magnification. In such a system a single ray of light passes through each lens and is consequently refracted and redirected to a particular point. When multiple rays of light are passed through the system, the refraction process bends each ray to converge on a single point that we can thus identify as the *focal point*.

This single point in simple terms is the part of the image that we can observe in focus, while the rest of the image will appear out of focus. Under higher degrees of magnification the area that can appear under focus will vary greatly, and be highly conditioned by the natural topography of the surface under observation. By adjusting the distance of the lenses we can manually oscillate the location of the focal point, focusing in on different areas of the image. The area where the rays of light converge and the degree of the image that can be focused on is considered the *Depth of Field*. The additional consideration of where the surface has to be positioned in relation to the lens is a measurement of distance called the *working distance*. Finally the surface area of the object that can be visualised under different degrees of magnification is considered the *Field of View* (FOV). In order to assess and evaluate the nature of a microscope and its efficiency under different circumstances, we should assess each of these parameters and find the means of observation that provides the best results (Bradbury and Bracegirdle, 1998).

In recent years, with the inclusion of digital cameras, automated focusing systems and computerised automation of image processing, new digital microscopes can digitally process images with different areas in focus to provide a digitally constructed image where the entire image is in focus. This is autofocus function is now included in most modern microscopes. In many cases this can be used to produce 3D digital reconstructions of these surfaces through a series of computed calculations and depth synthesis functions.

These new advances in 3D digital microscopy have proved particularly useful when measuring and evaluating the precise topographic nature of surfaces, proving an especially interesting advance when considering the study of microscopic variations of bone cortical surfaces. Perhaps the most significant advance in optical magnification, however, can be seen in the use of confocal microscopy. Standard optical microscopy presents the distinct disadvantage of the area of the image that is in focus when carrying out analytical studies. Confocal microscopy employs the use of a spatial pinhole that ensures the exclusion of "out of focus" light in the process of image formation. This fine ray of light, also known as a *laser,* is able to

produce images of a finer optical resolution. With the eventual inclusion of different types of light in the visible spectrum, such as fluorescent light, the quality of images are considered much greater with a quality of 3D imaging that is much closer to reality than images "artificially" produced using image reconstruction in digital microscopy.

All of these techniques present distinct advances that have been noteworthy in fields such as biology and engineering. Digital systems and technological advances have also been able to provide quantitative means of studying microscopic surfaces, producing a complementary development that has been of great value in scientific research. Archaeological applications are numerous and have been of great interest when considering BSM analysis as well as traceological use-wear studies (Stemp et al., 2015; Fernández-Marchena et al., 2016; Martín-Viveros, 2016; Mariciani et al., 2018; Wierer et al., 2018).

BSMs in general are minute variations in osteological materials, requiring at least 40x magnification in many cases for a superficial analysis of their properties, context and precise location. Common taphonomic practice highlights how identification of BSMs should never be performed solely on diagnosis through the naked-eye (Blumenschine et al. 1996). While there exists the possibility of "quasi-macroscopic" identification of BSMs, confirmation of a classification or taphonomic diagnosis is generally performed with the use of a handheld loupe with at least 10 or 20x magnification. More detailed observations and analytical studies in general, however, require magnifications of up to 40x and above.

The distinct benefits that microscopy provide for taphonomy can clearly be seen in a more detailed evaluation of BSM features and the precise nature of these qualitative features. With the additional use of metric tools and digital processing of microscopic measurements, in some cases with an accuracy and resolution of up to nanometres, microscopy provides a highly accurate and objective means of processing the microtopography of osteological surfaces.

In the 1980's a great deal of studies consisted in the use of Scanning Electron Microscopes (SEM), especially providing benefits for the visual perception of texture and the associated resolution provided (Potts and Shipman, 1981; Shipman and Rose, 1983; Shipman et al., 1984a, b; Andrews and Cook, 1985; Behrensmeyer et al., 1986; Cook, 1986; Olsen, 1988; Olsen and Shipman, 1988). Confocal microscopy on the other hand has also provided a significant development for microtopographical studies (Archer and Braun, 2013; Pante et al., 2017a; Otárolla-Castillo et al., 2017), alongside high resolution optical microscopy (Bello and Soligo, 2008; Bello et al., 2009, 2016). Combinations of these techniques have even proven to provide a taphonomic perspective of archaeological artefacts from the "inside-out" (Bello and Galway-Witham, 2019).

Nevertheless, microscopy cannot be considered the only means of digitally visualising BSMs with high resolution techniques. The use of photogrammetry and laser scanning devices have proved to be a key component for recent studies into BSM metric analysis. Photogrammetry consists in the computed reconstruction of objects using a large quantity of multiple photos from different angles using a fine quality camera. Using these techniques, archaeologists have been able to virtually reconstruct everything from entire archaeological assemblages to microscopic features of anomalies in bone.

Additionally, the introduction of laser and structured light scanning, have been able to reach new levels of high resolution at a lower cost while remaining more portable (Maté-González et al., 2017b). These pieces of equipment have also been useful in the development of new

methodological approaches in cut mark analysis (Courtenay et al., 2017).

An important consideration amongst all these different approaches, however, requires a reflection on the advantages and limitations provided by each piece of equipment. For example, microphotogrammetry and laser scanning have proved to be a much more accurate 3D means of representing archaeological artefacts, however their ability to detect minute variations in BSMs and more superficial traces have failed on multiple accounts. Future investigation must lie heavily on our means of obtaining data, especially traces and anomalies that present minute differences that consequently produce higher degrees of equifinality. Nevertheless, a further point to be developed is the means in which all this data is processed, highlighting the need for more robust and efficient statistical models when confronting the large quantity of high resolution data we are currently producing.

2.8. Geometric Morphometrics

2.8.1. An Introduction to Geometric Morphometrics

Geometry can be defined as a branch of mathematics that confronts the properties and relations of points, lines, surfaces, solids and other high dimensional analogues (Oxford Dictionary, 2018). This field can be additionally includes questions regarding properties of space, shape, size and the relative position of figures. The word in itself comes from the Greek terms "*geo*" and "*metron*", respectively meaning "earth" and "measurement", yet can be seen to have been initially conceptualized in the 6th century BCE. The real definition of this can clearly be marked by the work of the Greek mathematician Euclid, 300 BCE, in the city of Alexandria, Egypt, with the publication of *Elements* (Bilingsley, 1570).

With the further development of statistics in the 5th century BCE, a more empirically organized form of processing, collecting, analyzing, presenting and interpreting data was perceived, providing a much more conventional means of confronting mathematical questions of probability. After serious development of statistics in the 18th century CE, the 19th century saw the emergence of a subfield called morphometrics, consisting in the study of measurement (*metron*) and shape (*morphe*) (Mitteroecker and Gunz, 2009). The introduction of Frances Galton's correlation coefficient provided a key paradigm shift (Galton, 1888), presenting a powerful tool for processing biometric data. By 1907, the additional presentation of quantification methods through two-point shape coordinates, known as Bookstein-shape coordinates (Bookstein, 1991), revolutionized the statistical evaluation of morphometric data (Galton, 1907), efficiently processed through multivariate statistical techniques.

The 1980's saw a crucial development in morphometric studies, produced by the invention of coordinate-based methods, deformation grids and the conceptualization of the statistical theory of shape (Kendall, 1981, 1989; O'Higgins and Johnson, 1988; Bookstein, 1989, 1991; Goodall and Mardia, 1993; Bookstein, 1998; Dryden and Mardia, 1998). The computer became a fundamental tool within this academic development, providing an efficient means of exploring and visualizing data that can be considered too complex for the human brain to single-handedly comprehend. From this point of view, it can be seen how global technological developments have provided an incredible quantity of methods that allow certain disciplines, from mathematics to science, to flourish and adapt; providing more powerful means of quickly processing vast amounts of

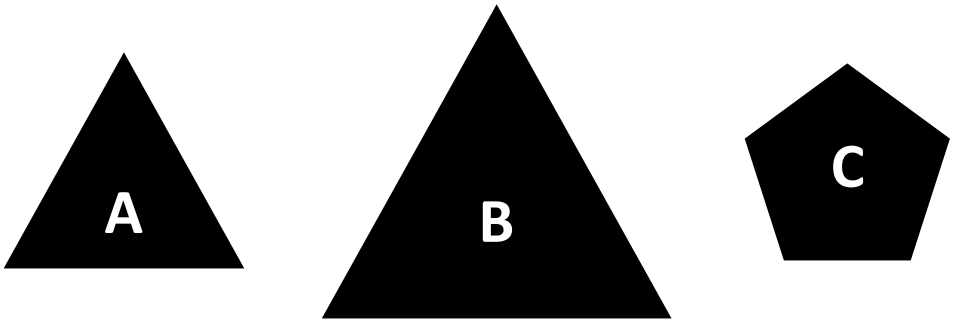


Fig. 2.10. - Figure presenting differences in shape and form. Triangles A and B are of the same shape yet have different forms, while pentagon C is of a different shape and form.

data. The development of statistical software, alongside high definition visualization and reconstruction techniques, considerably revolutionized these fields and have provided the backbone for incredibly advanced research that before would have been considered incomprehensible.

This new approach to processing shape data eventually coined the term *geometric morphometrics*. The value of this discipline can be defined through its ability to preserve the geometry of landmark configurations throughout analysis, thus permitting a valid representation of statistical means of processing *shape* and *form* (Mitteroecker and Gunz, 2009). As a product of this, an analyst is able to empirically and statistically quantify morphologies, removing subjectivity through qualitative classification of objects through simple observations.

A fundamental concept within morphological studies can be defined through the differentiation between *shape* and *form* (Oxnard, 1986; Goodall, 1991; Jungers et al., 1995; Mitteroecker and Gunz, 2009; Mitteroecker et al., 2013). Differentiating between the two is crucial in understanding and confronting questions of geometry. *Form* can be defined as the visual shape or configuration of the individual (Oxford Dictionary, 2018), however within geometry, form fundamentally includes the size of an object. In order to fully comprehend the morphology and configuration of certain individuals, we resort to analyzing *shape*. *Shape*, as opposed to *Form*, is the pure outline or geometric configuration of the individual, regardless of size. *Shape*'s definition importantly relies on describing the external form, contour or outline of the individual under observation (Oxford Dictionary, 2018). This consequently excludes the variables size, position and orientation. As a result of this, through comparison of both *Form* and *Shape*, we can truly assess the morphological differences between individuals (Oxnard, 1986; Jungers et al., 1995).

Fig. 2.10. presents an example of three individuals. A and B present the same shape yet different forms, while individual C presents both differences in shape *and* in form. In these cases, positioning points on the corners of each individual would be sufficient in fully describing their morphological characteristics and *shape*, regardless of their rotation and translation (Fig. 2.11.). If we are then to consider the distance between these points, we would be making observations regarding their *form*. The correct nomenclature to describe these points lies s in the word *landmark*.

A landmark can be defined by a "point of correspondence on each object that matches between and within populations" (Dryden and Mardia, 1998). Landmarks can be further divided into three main types; mathematical, scientific and pseudo. Mathematical landmarks are points located on an object according to a mathematical or

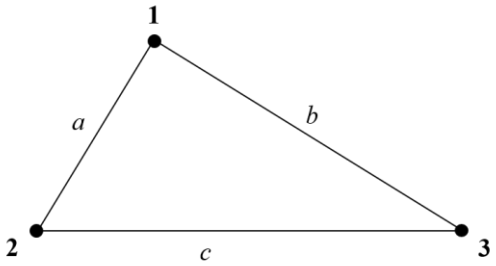


Figure 2.11. - Diagram explaining the theory behind shape and form variation, where landmarks 1 - 3 are sufficient in describing the pure morphological shape of an individual while distances *a*, *b* and *c* are required to describe its form

geometrical property of the figure. Scientific landmarks are defined by an expert, corresponding to a point that can be defined scientifically. In biology these can be anatomical features or points of biological derivation. Pseudo-landmarks are constructed points on the object that can be located in between scientific or mathematical landmarks. Further definitions of landmarks and their properties have led to the use of nomenclature such as Type I, II and III (Bookstein, 1991). According to Bookstein (1991), these types can be used to define particular biological points of interest. Type 1 landmarks occur at joins of tissues or bone, while Type II can be found in local areas of interest such as points of maximal curvature. Type III landmarks are external points, best compared with the aforementioned definition of pseudo-landmarks. These can be found at places of maximal diameter or centroids.

Another key concept in landmark data lies in their *geometric homology*. Landmarks are assumed to be homologous points located on all specimens that can later be compared and calculated in the form of *Cartesian coordinates* (O'Higgins and Johnson, 1988; Bookstein, 1989; Goodall, 1991; Hall, 2003; Manríquez et al., 2006; Klingenberg, 2008).

A further type of landmark can be seen in the semi-landmark. A semi-landmark is simply defined as a point located on a curve (Bookstein 1996; Green 1996; Perez et al., 2006). Semilandmarks require calculation in order to be placed, treating these points as missing data that are digitally computed and estimated in order to minimize net *bending energy* (Gunz et al., 2004a). The development of semi-landmark models were originally constructed to confront the 'strictness' of landmark selection criteria, especially concerning the fact that in many cases explicitly targeting a point can overcomplicate a study. This is especially evident when the target point for a landmark is not there, a frequent issue in archaeology, for example (Gunz et al., 2004b; Gunz et al., 2009; Arbour and Brown, 2014).

In order to extract *shape variation* from a configuration of landmarks, a statistical process of scaling, rotating, translating and superimposing landmark configurations is carried out. This systematically produces a standardized computed construction ready for analysis. Through this, concepts of size and orientation of points are removed.

One of the main methods in confronting this question can be found in the *two-point shape coordinate* approach (Bookstein, 1991). With the introduction of more than two dimensions, a more generic and common means of computing shape coordinates is the *Procrustes Superimposition* method (Fig. 2.12.), and in cases where more than two forms are present; *Generalized Procrustes Analysis (GPA)*. Procrustes superimposition consists in a least-squares approach whereby each landmark configuration is;

1. Translated so they share the same *centroid*. This is done by calculating the mean of all points (its *centroid*) (Bookstein, 1991).
2. Scaled so that they all share the same *centroid size*. This is done through calculating the root mean square distance from the points to their translated origin.
3. Superimposed so that the sum of the squared *Euclidian distances* is

minimal (Rolf and Slice, 1990; Lele and Richtsmeier, 1991, 2001).

GPA, as opposed to simple Procrustes superimposition, relies on a specific algorithm where the resulting coordinates are averaged producing a reference orientation different to that of Procrustes superimposition (Gower, 1975). Figure 2.13. provides a visual representation of this process, using the landmark data from Courtenay et al. (2017).

Visualisation of landmark data either in 2D or 3D can be described by *shape space*, where a single point can be used to represent the shape of each individual. *Kendall shape space* describes a sphere, rather than a flat plane, where each point can be plotted (Kendall, 1981). When considering more complex shapes and forms in 3D, the calculations and dimensions used to describe this sphere become more developed (Kendall, 1984; Monteiro et al., 2000). As described by Kendall, if the distribution of individuals in shape space are independently and identically distributed, then shape can be described as isotropic. Variations present in this isotropy can then be used to differentiate between samples (Bookstein, 1991; Dryden and Mardia, 1998). Patterns in variation and covariation are thus statistically differentiable (Rohlf, 1999; Slice, 2001).

In reality, the degree of variation and quantity of variables included in a study of this type can never empirically reach a degree of linearity (Mitteroecker, 2007; Mitteroecker and Gunz, 2009). The curvature of shape space is a large conditioning factor in this observation. Statistical linearity can only be established through constructing a tangent to the reference mean shape (Fig. 2.14.). Euclidean distances in *tangent space* closely resemble Kendall's shape space, and can thus be used in preparation for multivariate statistical analysis. Preparation, construction and calculation of tangent space can be performed based on the geometric properties of orthogonal projections. This is done through projecting each point onto the space perpendicular to the vector of shape coordinates of the reference shape (Rohlf, 1999; Monteiro et al., 2000). Common malpractice of geometric morphometric analysis tend to ignore this factor (Mitteroecker and Gunz, 2009; Mitteroecker, 2007).

Procrustes form space does not differ too much from the previous explanations of shape space. The main difference, as previously defined, lies in the variable, *size*. Form space can thus be calculated including the natural logarithm of centroid size (Mitteroecker et al., 2004, 2005; Gerber et al., 2007).

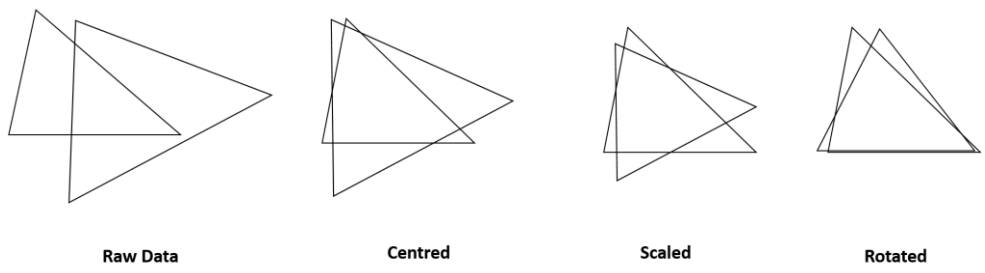


Figure 2.12. - Example of some of the different processes involved in Procrustes Superimpositions.

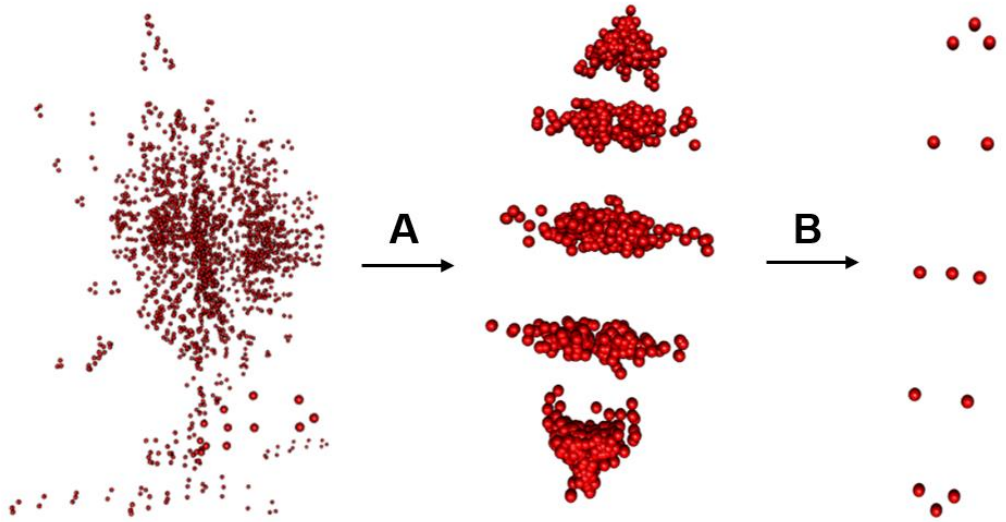


Figure 2.13. - Example of the transformation of raw landmark data into a landmark model. First beginning with all the raw landmark coordinates, followed by (A) the process of General Procrustes Analysis, using a mixture of scaling, rotation, translation and superimposition. Finally (B) the calculation of mean shape. Landmark data used to create this figure is from Courtenay et al. (2017)'s established 13 landmark 3D model for the study of cut marks.

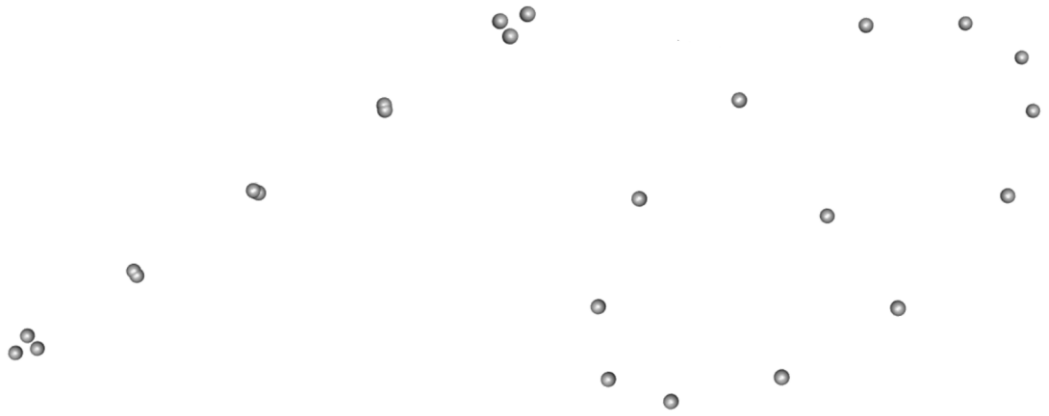


Figure 2.14. – Example of two different extremes in shape variation. Landmark data from Courtenay et al. (2017).

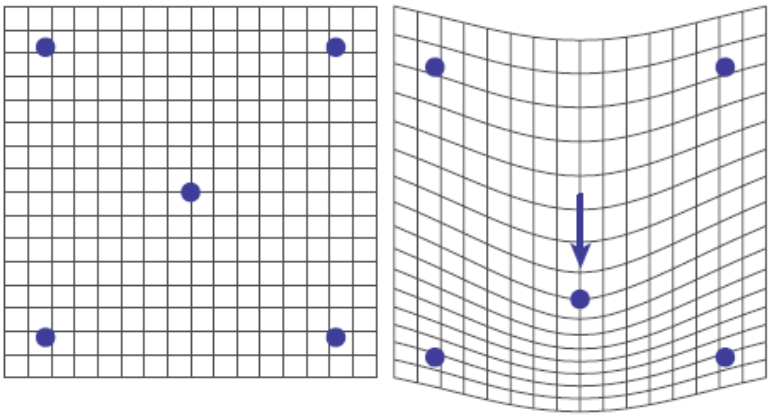


Figure 2.15. – Example of TPS deformation grids with 5 landmarks projected against the template configuration (left) and the target configuration (right). Figure from Mitteroecker and Gunz (2009)

Visualising changes and differences in shape can be a fundamental tool when beginning to compare morphologies. For this we come across the use of *Deformation Grids*. This technique is based on mapping out the homologous points in a visual manner. The first attempts at applying this concept were carried out by hand (Thompson, 1915, 1917), however with the use of computers, fast algorithms are able to calculate these grids through the use of *Thin-Plate Splines* (TPS) (Bookstein, 1989, 1991). TPS are calculated by mapping out point configurations while the space in-between points is 'smoothly interpolated' (Figure 2.15.). Smoothness is fundamentally defined by minimizing the bending energy of a deformation, calculated by the squared second derivate of the deformation. The basic TPS template consists in a grid where points are projected. Differences in point configurations then warp the grid and aid in visualizing differences between shapes.

Combined with advanced statistics, geometric morphometrics provide a powerful tool that can be used to quantify data that in essence are mostly qualitative and visual. Through this we can assess morphological differences, variation and even come to classify objects based on their physical appearance, shape and form.

2.8.2. Applications in Archaeology

The arrival of geometric morphometrics (GMM) in Archaeology have facilitated a major advance in our current understanding of the paleoanthropological record. Some of the most significant methodological advances in fossil classification have been fruit of GMM analysis of cranial morphology. Through this, it can be seen how GMM have become a fundamental tool in the study of human evolution.

The "morphometric revolution" in archaeologically can arguably be dated back to the 1990's (Rohlf, 1993), however this field is still relatively new and is under a constant state of development (Baab et al., 2012). Applications of GMM in human evolution can be described under three main objectives; fossil classification, morphological differentiation and the comparison of variables.

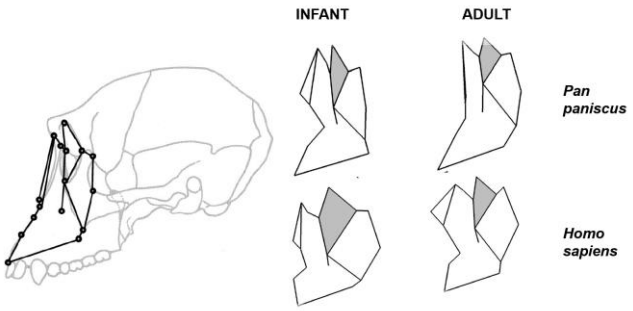


Figure 2.16. – Differences in Facial Ontogenetic Shape Trajectory Variation between bonobos and humans. Figure adapted from Cobb and O'Higgins (2004)

Initially, studies of this particular nature utilized scientific landmarks, mostly Type I in nature, relying on anatomical features for precise targeting (Richtsmeier et al., 1995; Dryden and Mardia, 1998; Adams, 1999; Oxnard and O'Higgins, 2009). Developed use of sliding semi-landmarks utilizing their ability to quantitatively capture curvature, however, have provided a key advance in the field (Bookstein 1996; Green 1996). This advance has been able to generate significant results in palaeoanthropological GMM (Gunz et al., 2005; Perez et al., 2006).

Part of studies in variation focus greatly on ontogenetic and evolutionary changes (Rice, 1997). Functional morphological studies, phylogeny and hominin systematics, however, also play a vital role in evolutionary GMM. All of these studies into variation and covariation of morphological features are a fundamental preliminary stage before classification studies can begin. Understanding the true nature of ontogeny (Rice, 1997; Bastir and Rosas, 2004a, b; Carlson, 2005; Bastir et al., 2006, 2007), for example, becomes vital when trying to classify the species of an infant individual (Gunz et al., 2011; Gunz and Bulygina, 2012).

Understanding ontogenetic variation and development in our evolutionary history has produced a multitude of different revelations about our development as a species (Kondo et al., 2000; Lieberman et al., 2002; Cobb and O'Higgins, 2004; Kondo et al., 2005; Zollikofer and Ponce de Léon, 2004, 2010; Rosas et al., 2017). A great deal of these studies in ontogenetics uses GMM to estimate allometry, providing the necessary tools to study growth trajectory trends between different species (Klingenberg, 1998). Allometry is based on shape variation induced by general size variation. Using an estimation on the correlation of shape with increase in size over time, for example, makes differentiable trends amongst certain species more evident (Fig. 2.16.) (Cobb and O'Higgins, 2004). Studies of this type are able to highlight intra-specific characteristics early in ontogenetic development, using extant and extinct primate species to provide stronger insights into developmental patterns (Frost et al., 2003; Ackermann, 2002; Cobb and O'Higgins, 2004). Some studies have even been able to link results in studying craniofacial growth trajectory to phylogenetic relationships (Lieberman et al., 2002; Frost et al., 2003; Harvati et al., 2010).

One major issue that becomes increasingly clear with paleoanthropological studies of this nature is the scarcity and survival rate of fossils in the paleontological and archaeological register (Rohlf. 2000; Gunz et al., 2004b). This issue provides an alarming constraint in the quality of results drawn from most studies, especially considering the statistical significance of such small sample sizes (Chartier and Allarie, 2007; Faul et al., 2007; Reinhart,

2015). While recent advances in developed simulations have been able to overcome some issues in ontogenetic studies (Freidline et al., 2012, Gunz et al., 2012; Freidline et al., 2013), the most significant advances in GMM applied to the incomplete fossil record have been found in computed reconstructions of fossils (Gunz et al., 2004b, Benazzi et al., 2009; Gunz et al., 2009, 2011, 2012; Arbour and Brown, 2014).

The work of Gunz et al. (2004b; 2009) have revolutionized GMM research through the use of statistical algorithms and calculations to predict the location of missing landmark data. The Expectation Maximisation (EM) algorithm (Dempster et al., 1977) can be used to handle missing data through the calculation of likelihood probabilities which in turn are used to predict the exact location of missing landmarks. Through the random knockout of landmark data on complete skulls, Gunz et al. (2004b) were able to compare three statistical models in missing data handling; Statistical Regression, Geometric Reconstruction and Mean Substitution. Their results were able to highlight mean substitution (a relatively simple procedure common in social sciences) as the method that produced the highest margin of error, while statistical regressions were the best means of calculating landmark positioning. Geometric Reconstruction employs the use of TPS to calculate the missing data based on existent complete individuals while statistical regression calculates an estimated probability of where the missing landmark should lie.

Further development of these results made the most of natural 'symmetry' in crania to calculate the positioning of missing landmarks (Gunz et al., 2009). More interestingly, however, were the results when applying these symmetrical TPS based methods to the correction of taphonomic distortion in individuals such as Arago XXI ('Tautavel Man'), from Caune de l'Arago (Fig. 2.17.). The interesting case of Arago XXI presents taphonomic post-depositional diagenetic processes that appear to have sheared and deformed the skull of a pre *Homo neanderthalensis / Homo heidelbergensis* individual (de Lumley and de Lumley, 1971; Seidler, 1997) Due to the fairly uniform distortion of the skull, however, TPS warping could correct the shearing and rectify the cranial morphology of this particular individual (Gunz et al., 2009). This could be particularly useful when applying GMM classification tests to diagnose the Arago XXI individual as either Neanderthal, pre-Neanderthal (Grimaud-Hervé, 1998), or better defined directly as *Homo heidelbergensis* (Guipert et al., 2004, 2014). Diagenetic modifications would make this particularly difficult without the un-shearing of the skull, especially when considering the already existent mixed cranial features of this particular individual (de Lumley and de Lumley, 1971, 1973; Day, 1986; Cela Conde and Ayala, 2013).

Classification of hominin fossils using GMM have provided a much more quantitative and objective approach to species differentiation (Fig. 2.18.). In paleoanthropology, many authors argue that the combined use of form and shape are powerful tools in classification, while form provides the highest degree of classification in hominin fossils (Mitteroecker and Gunz, 2009; Mitteroecker and Bookstein, 2011; Mitteroecker et al., 2013).

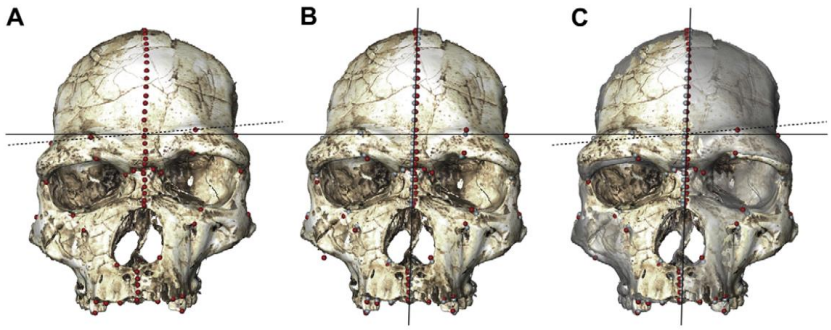


Figure 2.17. - The GMM digital correction of taphonomic distortion in the skull of Arago XXI ('Tautavel Man'). (A) Before correction. (B) Correction. (C) Comparison between the two. Source: Gunz et al. (2009)

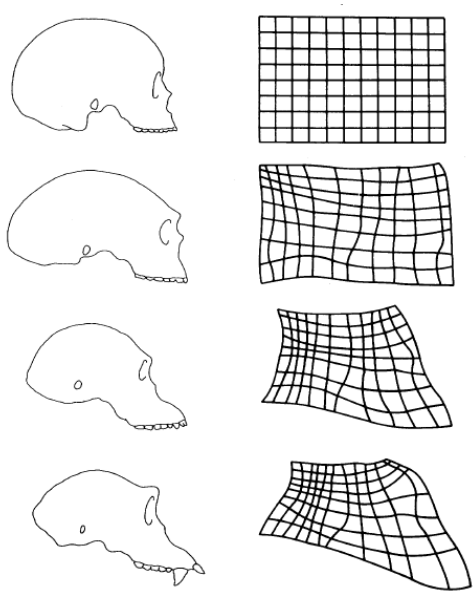


Figure 2.18. - Different TPS representing shape variation in different skulls from different species. Source: Oxnard (1986)

In order to understand the classification process, studies have looked deeply into using GMM to describe variations in morphological characteristics of hominin fossils (Baab et al., 2012; Mitteroecker et al., 2012). Through analyzing variation in hominin cranial morphologies, many analysts have been able to refine the classification process using GMM (Klingenber and Marugán-Lobón, 2013), and thus create a reference sample with which to compare fossils (O'Higgins and Johnson, 1988; Bookstein et al., 1999; Bruner et al., 2004; Bromage, 2008; Gunz et al., 2010, 2012; Artheya, 2012; Bruner et al., 2013). Combined with the use of multiple statistical tests, GMM is used to compare the significance in differences between samples, testing to find which sample is closest in morphology (Gunz et al., 2009, 2011, Freidline et al., 2012; Gunz et al., 2012; Gunz and Bulygina, 2012).

The power of GMM classification tests has been debated for quite some time. While most GMM statistics use advanced algorithms and testing, in many cases these tests may fall short (Albrecht, 1992; Jungers et al., 1995; Arnqvist and Mártensson, 1998; Rohlf, 2000). Classification tests, however,

will be dealt with in great detail in future sections of this paper.

Physical anthropology and paleocraniology have had a particularly strong protagonism in archaeological GMM. Its wide use and interesting results, however, have drawn attention to the advantage that advanced statistical tests and methods can present for other fields of prehistoric research and human evolution. The development of GMM applied to other fields of human evolution however, emerged at a much later stage. In many fields of classification, the overwhelmingly subjective and qualitative features of typological analysis are based heavily on experience and intuition in traditional approaches to prehistoric studies (Dunnel, 1971; Gero and Mazullo, 1984; Hiscock, 2001). In lithic studies however, the use of biometric methods alongside multivariate approaches eventually began to overcome typological subjectivity (Laplace, 1972; Roe, 1981; Thompson, 1981; Carbonell et al., 1983; Laplace and Livache, 1986; Bordes, 1988; Wynn and Tierson, 1990; Carbonell et al., 1992; Crompton and Growlett, 1993; Débenath and Dibble, 1994; Merino, 1994; Rodríguez-Álvarez, 1997; Ollé, 2003; Vergés, 2003; Franco et al., 2005; Grossman et al., 2008). Even if only slightly in some cases.

The first use of lithic GMM can be found in the landmark based methods for the study of bifacial handaxes (Brande and Saragusti, 1996). Similar studies can be found in Lycett et al. (2006) who applied these methods to three dimensional studies of Pleistocene cores.

These initial advances made an interesting first step in the right direction for lithic GMM studies, however took a while to truly develop. More substantial practices in GMM lithic studies were orientated around stylistic and functional properties of Upper Palaeolithic and Holocene projectile points (Cardillo, 2006; Castiñeira et al., 2009; Cardillo, 2010), combined with microwear analysis (Franco et al., 2009), to truly understand the functionality between different point morphologies. Their impact in the field, however, were mostly limited to publications in low impact journals in the Spanish language, limiting their methodological re-productivity in other countries. Regardless, their importance can also be seen to have eventually reached an application in the study of some scrapers (Cardillo and Charlin, 2009; Cardillo, 2009).

Further notable GMM approaches to the study of Acheulean handaxe technology (Fig. 2.19.) can be found in Costa et al. (2010), Iovita et al., (2017) and García-Medrano et al. (2018).

Needless to say, the impact of GMM in modern archaeological studies has been fairly limited, despite the potential this field may have on the future of quantitative and objective typological analysis.

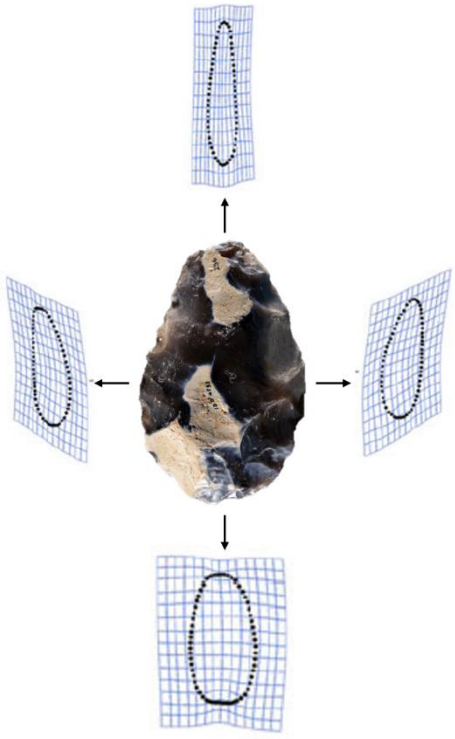


Figure 2.19. – Acheulean bifacial handaxe planfrom morphological variation from Boxgrove (Sussex, UK). Figure adapted from García-Medrano et al. (2018)

2.8.3. Applications in Taphonomy

The first approximations to morphological studies in taphonomy can be found in the work of Walker and Long (1977), presented through their study of cut mark profiles (Fig. 2.20.). This paper was the first to make an attempt at understanding morphology, however, their work consisted in purely qualitative observations with no statistical backing behind their observations. After this initial reflexion on the V shaped cross section of a cut mark, morphological studies of BSMs remained scarce up until the beginning of the 21st century.

The biggest advance in morphological studies of taphonomic traces came about with the work of Silvia Bello (Bello and Soligo, 2008; Bello et al., 2009; Bello, 2011; Bello et al., 2013) who suggested a series of 6 measurements to best describe cut mark morphology. Bello's original parameters (Bello and Soligo, 2008) consisted in:

1. Slope angles
2. Opening Angle of cut mark
3. Bisector angle
4. Shoulder heights
5. Floor radius
6. Depth of Cut

These 7 metric variables (Fig. 2.21.) were able to succesfully differentiate between different tools (Bello et al., 2009), provide a successful application to the archaeological record at Boxgrove and Gough's Cave (Bello, 2011), and demonstrate the further potential these studies may have when applied to protable art (Bello et al., 2013).

Bello's interesting work provided the first quantitative approximation to the morphological study of cut marks, combining the use of complex statistics and 3D microscopy. Her work can also be classed as the first real pioneering study in BSM morphologies, considering the value of her results compared with those of Walker and Long (1977).

Bello's work inspired more authors to apply similar methodological approaches in cut mark studies with the use of different microscopic approaches (Boschin and Crezzini, 2011; Moretti et al., 2015; Braun et al., 2016; Fuentes-Sánchez et al., 2017; Duches et al., 2018; Stinnesbeck et al., 2018). The statistical power of multivariate statistics on biometric variables in these cases, however, present some limitations.

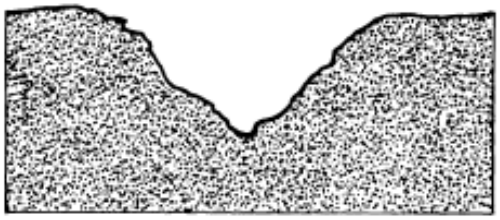


Figure 2.20. - The first approximation to cut mark morphology published in Walker and Long (1977)

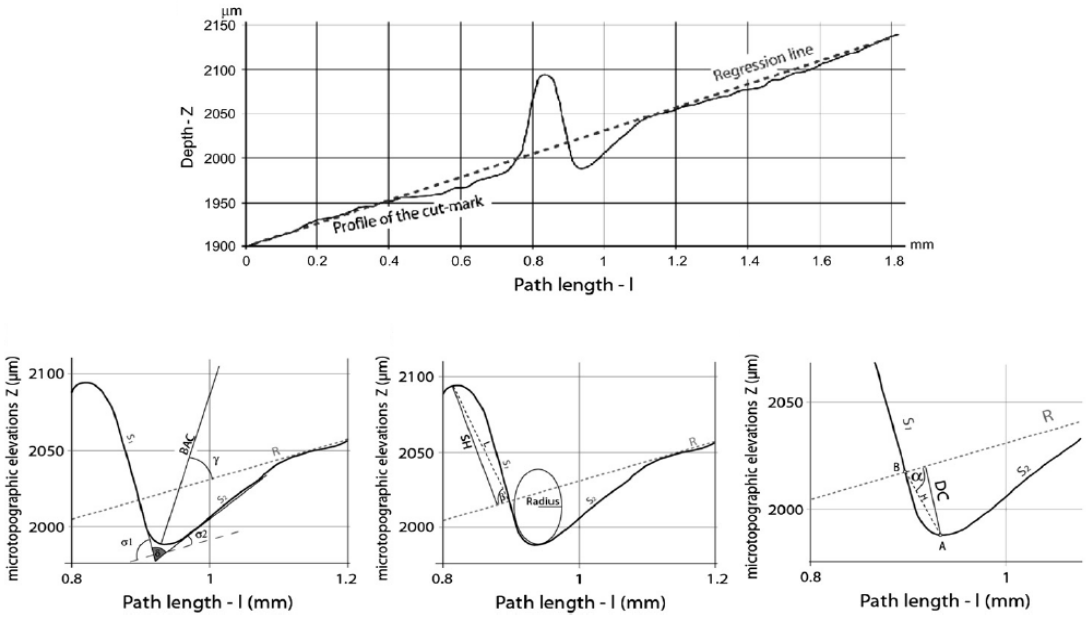


Figure 2.21. - The original description of 7 biometric measurments that permit the quantification of cut mark cross section profiles (Bello and Soligo, 2008)

The conversion of Bello's methods into a GMM approach came about with Maté-González (2015)'s first GMM analysis of cut mark profiles.

Employing the use of micro-photogrammetry, the aforementioned study converted Bello and Soligo (2008)'s biometric approach into a 2D 7-landmark model, combining 8 measurements and 7 landmarks to morphologically compare cross-section profiles. The 8 measurements described by Maté-González et al. (2015) are as follows:

1. Width of the Incisions at the Surface (WIS).
2. Width of the Incision at the Mean (WIM).
3. Width of the Incision at its bottom (WIB).
4. Opening Angle of the Incision (OA).
5. Depth of the Incision (D).
6. Left Depth of the Incision Convergent (LDC).
7. Right Depth of the Incision Convergent (RDC).
8. Angle of the Tool Impact (ATI).

Asides from the introduction of GMM into cut mark morphological analysis, Maté-González et al. (2015) began by assessing the significance of cut mark profiles across the entire mark, concluding that the most representive cross section profiles can be found between 30 and 70% of the mark's total length.

This pioneering use of taphonomic GMM opened multiple doorways, confronting studies according to cut marks produced by different raw material (Maté-González et al., 2017b), tool types (Yravedra et al., 2017a) and an application to the archaeological register (Yravedra et al., 2017a, b). Cut marks and other anthropic modifications have not been the only protagonists in this field, however, with multiple works also confronting the morphology of carnivore tooth scores in a similar fashion (Arriaza et al., 2017; Yravedra et al., 2017c).

These pioneering papers, however, worked using a combination of 2D and 3D methods, using 3D digitally reconstructed

models to produce 2D images that could then be used for GMM processing. The introduction of 3D GMM in taphonomy came about with two papers; Courtenay et al. (2017), for the study of cut marks, and Aramendi et al. (2017a), for the study of tooth pits. Courtenay et al. (2017)'s 13 landmark model (Fig. 2.22.B.) provided a higher degree of resolution for discerning raw materials and tool types from entire cut mark morphologies, while Aramendi et al. (2017a)'s 17-landmark model (Fig. 2.23.) provided a powerful tool for discerning carnivore agency using tooth pits.

In archaeological studies, the subsequent statistical combination of both 2D and 3D models have provided a key advance in the study of anthropic taphonomic trace morphologies (Courtenay et al., 2018a) (Fig. 2.22.A.). A combined effort in differentiating between carnivore produced tooth pits and anthropic percussion pits has also been presented using the previously described 3D 17-landmark model (Yravedra et al., 2018).

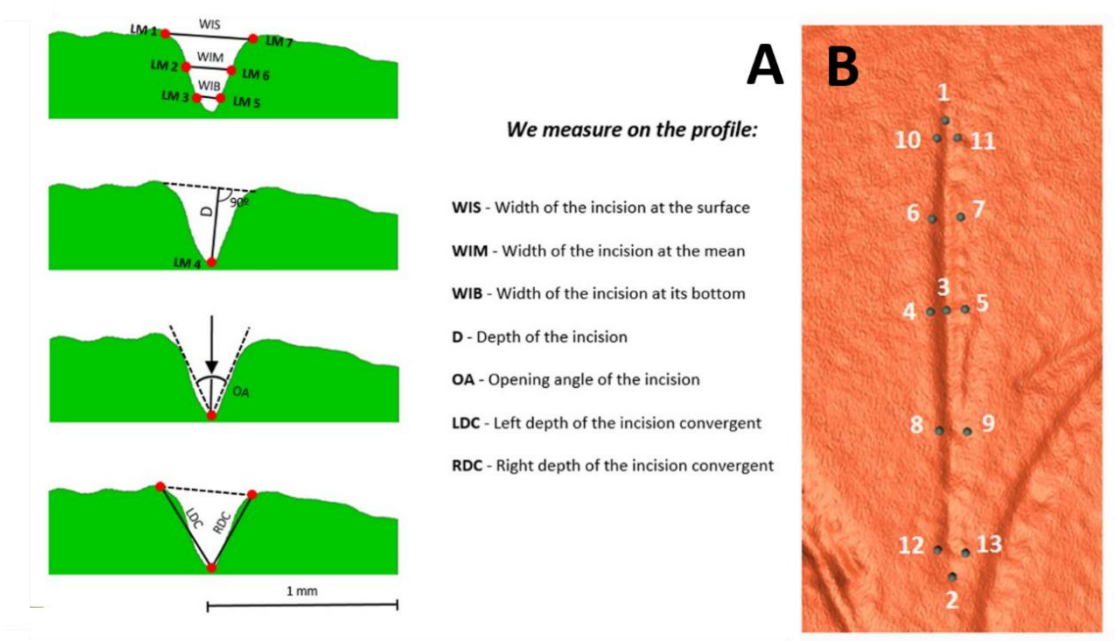


Figure 2.22. - Diagram presenting (A) Maté-González et al. (2015)'s 2D 7-Landmark Model, the adapted 7 measurements according to Bello and Soligo (2008) and (B) the 3D 13-landmark model according to Courtenay et al. (2017). The combined use of all models has been discussed in Courtenay et al. (2018b).

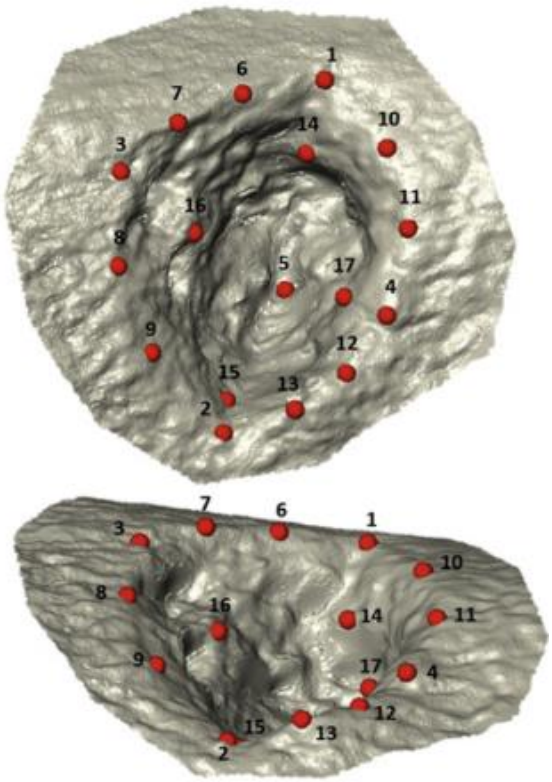


Figure 2.23. - Aramendi et al. (2017a)'s 3D 17-Landmark model for the morphological study of carnivore tooth pits

The development of GMM methods for the analysis of microscopic traces, however, have not come without their difficulties. From one point of view, the methodological approach required to digitally reconstruct 3D models of such microscopic features requires a sufficiently high degree of resolution. One objective behind many papers has included the comparison of 3D reconstruction methods, using 3D digital microscopy (Maté-González et al., 2017a, c), microphotogrammetry (Maté-González et al., 2015), and more recently structured light laser scanners (Courtenay et al., 2017; Maté-González et al., 2017a). Beyond this, the statistical treatment of data has also been a topic of development and debate. Work into the production of statistical packages for GMM studies (Adams and Otárolla-Castillo, 2013; Bonhomme et al., 2014; Palomeque-González et al., 2017; Adams et al., 2017), statistical software (O'Higgins and Jones, 2006; Klingenberg, 2011), as well as debates between Bayesian and Fisherian methods of data handling (Otarólla-Castillo et al., 2017; Courtenay et al., 2018a), have produced a great deal of interesting results. The statistical weight of biometric variables such as OA and ATI (Fig. 2.22.A.), for example, have been proven to be problematic in some studies, to the point where ATI is excluded from most current testings. OA, on the other hand, has proven to negatively affect some results (Maté-González et al., 2017a; Courtenay et al., 2018a), resulting in most cases requiring double testing; firstly including OA and secondly excluding OA. Needless to say, both the topics of microscopy and statistics are key components of this paper and will be discussed in great detail throughout.

Much like the use of variation and covariation studies in paleoanthropology, a

great deal of studies in taphonomic GMM have been dedicated towards variations in cut marks according to a number of different variables. While the different carnivores producing tooth marks is an evident variable to consider (Arriaza et al., 2017; Aramendi et al., 2017a; Yravedra et al., 2017c, 2018), in the case of cut marks, multiple variables are yet to be tested. The most important variables that have already been studied consist in raw materials (Maté-González et al., 2016, 2017a; Courtenay et al., 2017; Yravedra, 2017b), their geological granular composition (Courtenay et al., Under Review), tool type (Courtenay et al., 2017; Yravedra et al., 2017a), cutting angle (Otárolla-Castillo et al., 2017; Courtenay et al., 2018a) and bone density (Maté-González et al., 2018). Through our current understanding of the taphonomic register, however, some authors highlight a great deal of variables that are yet to be studied in detail. Among these, taphonomic alterations of BSMs and cortical surfaces are a key component (Pineda et al., 2014, 2019). Equifinality with other taphonomic traces such as trampling marks also requires further testing (Pante et al., 2017a; Gümrükçu and Pante, 2018; Orlikoff et al., 2018). Other factors that have been studied and mentioned, but are yet to be tested using GMM means, are the sharpness of the cutting edge (Braun et al., 2008, 2016), as well as tests with other raw materials such as bamboo (Spennemann, 1990; West and Louys, 2007; Bonney, 2014) and shells (Choi and Driwantoro, 2007; Weston et al., 2015).

Needless to say, GMM methods and their application to the study of taphonomic BSMs in archaeological sites have begun to reveal incredibly interesting results (Aramendi et al., 2017a; Linares-Matás et al., 2017; Yravedra et al., 2017a, b; Yravedra et al., 2019; Courtenay et al., Under Review).

Among these studies, reflections on ancient hominin behaviour and cognitive complexity have also become possible (Rodríguez-Hidalgo et al., 2018; Courtenay et al., Under Review; Rodríguez-Hidalgo et al., Under Review).

2.9. Artificial Intelligence

The earliest methods of recording data required the work of a human being to first observe, and second record the observation. These techniques mostly resolved around astronomy and observational ponderings of our environment, however, with an increase in scientific knowledge, databases are used for pretty much every aspect of our daily lives. Technology, as presented in many aspects of this paper, has generated more systematic means of documenting observations and quantitatively recording data, however, technology in recent years has even come to a point where systematically automated recording process are more commonplace in an ever-growing computer friendly society.

Computer sensors in recent years have attributed to an increased richness in scientific research, being able to capture basic human senses such as seeing and hearing. This approach to scientific research is an incredible advance that has a powerful use in fields such as biology and pharmacy, where more subjective means of recording data could be the difference between life and death. It is important, however, to consider that while sensor recorded data is a means of reducing subjectivity, we cannot use them to present absolute truth (Lantz, 2013). A black and white image reveals different information to one taken in colour. A photo taken using Scanning Electron Microscopes will reveal different

details to observations made using a digital microscope.

Larger and more complex means of collecting, organising and recording data have required the processing capacity of machines, however with this advance has come the need for a systematic means for processing it all. The field of study that confronts the development of computer algorithms for transforming data in this way is known as Machine Learning (ML), while a further subdiscipline of this is known as Deep Learning (DL).

ML is a subdiscipline of the much larger discipline of Artificial Intelligence (AI), within the scientific field of computer science. AI can be considered a means of preparing devices to correctly interpret external data, taking decisions and consequent actions that optimise the chances of achieving a goal (Nilsson, 1998; Poole et al., 1998; Russell and Norvig, 2003; Legg and Hutter, 2007; Kaplan and Haenlein, 2018). However the field of AI much like any other discipline is composed of much smaller components that work towards the investigation of the same goal.

Founded in 1956, AI's initial conception and founding was orientated towards the learning and solving of different mathematical problems in algebra (Crevier, 1993). With significant developments unfolding produced by the students of AI's initial founders, research moved on to the trivial use of computer models to analyse chess game strategies (Samuel, 1959). Historiographically fuelled by the cold war, it was not long before AI was implemented into major government military projects (Crevier, 1993; McCorduck, 2004), and in business marketing schemes (Leprince-Ringuet, 2018).

While AI over the years has come to be associated with futuristic concepts of machines learning how to train themselves, resulting in science fiction themed images of computers taking over the world, the reality is far from this and can currently be considered more like training an employee (Lantz, 2013). The machine is taught how to process the data, and is then used to transform and process information with a resulting intelligent output.

The fundamental need for this in science has grown from an alarming increase in variables, data and observations that the human mind is unable to process efficiently. With this increase has arisen a need for additional computing power, thus requiring more statistical methods for confronting data analysis. Part of the main statistical advances in this field is known as *data mining*. This method consists in the 'mining' or large databases in order to provide a more systematic means of managing large chunks of data. ML algorithms are a useful tool in carrying out more complex data mining. ML algorithms have proven to be the most successful approaches to identifying patterns in data samples, reaching ridiculously high classification rates.

ML in a nutshell is the means of teaching a machine to learn if it is able to take experience, and utilize it in order to improve its performance on similar experiences in the future (Mitchell, 1997). The basic learning process (in humans as much as machines) consists in data retrieval (the input), whether this be through memory, observations or experience, which is then translated into a broader representation of data (abstraction) with a final output where the abstract data is used to form the basis behind an action (generalization). ML spends time selectively managing smaller sets of key ideas in order to build an outline or concept. This can be used to define relationships amongst the present information, and in the process, is able to depict models that can be used to confront similar problems.

Learning algorithms confront raw data using structured techniques that assign meaning to the data (abstraction). Prior to abstraction, in reality the data is a simple string of numbers with no meaning. Simple "yes" and "no" concepts to a computer remain in binary concepts of 1s and 0s. The abstraction process is a technique used to construct logical structures that assist in turning this raw sensory information into meaningful insights (Lantz,

2013). This summarised representation of data is called a *model*. Models represent explicit descriptions of data through structured patterns. Models can thus exist in the form of equations, diagrams, cluster groupings and logical rules.

Under constructing the different models, the machine is not typically capable of choosing among models. The machine is capable of carrying out the learning task in order to construct the model, however input dictating which method for data abstraction is necessary is up to the analyst. The process in fitting a particular model to a dataset is what is known as *training*. Use of the term training is a powerful means of describing the actual process undertaken when the model is fitted to the data, taking into account the construction of models imposed by the human teacher to the machine student. In this sense, the most efficient means of ML is through the combination of multiple algorithms with the eventual evaluation of model performance (Kuhn and Johnson, 2013). It is important to highlight, therefore, that not all models are capable of presenting a paradigm shift that effects our knowledge of information. These algorithms are more likely to produce a revelation that helps plan future testing, whereby underlying relationships that have always been present are now made evident. The conceptualisation of this information under a new format creates the value for combined model usage.

The actual fitting of a model to data is performed using a series of mathematical concepts that employ a combination of linear algebra, probability theory, calculus and statistics. Through multiple mathematical equations, the models being built employ numerous calculations to arrive at an answer, using *weights* within these calculations to compute an output. The way these initial weights are calculated and updated requires different strategies depending on the problem-solving situation at hand. More in depth descriptions of these mathematical concepts, however, will be explained in detail in the methods section of this study.

The final process of learning is known as *generalization*. This concept confronts the application of abstract knowledge and its preparation in future actions. In general, the machine is capable of revealing infinite connections between data strings, however will be unable to utilize this information if no method of generalization is used to tune the abstract knowledge and thus determine what is useful and what is not. The process of tuning data models and pruning decision trees, for example, helps remove conditioning variables that generate more statistical noise than they help provide answers. This provides manageable data with more efficient processing means of identifying important findings. Thankfully, ML is capable of doing this efficiently without the need to examine each variable one by one. ML algorithms employ tuning functions that quickly divide the data, employing *heuristics* (educated guesses) to evaluate the importance of each variable. Heuristics can be neatly summarised by the human concept of instinct and our capability of making snap decisions. In ML, this can be seen in the speed at which probabilities and event likelihoods can be recalled. ML algorithms do this through estimation of likelihood ratios, considering statistical sensibility and sensitivity.

Sensitivity and sensibility can be explained through the ratio and proportion of Type I and Type II errors. This concept is fundamental in statistics when considering the calculation of confusion matrices and classification tables. This common practice in medical statistics defines the likelihood of correctly diagnosing an illness (sensitivity) or the likelihood of correctly identifying the healthiness of an individual (specificity) (Fawcett, 2006). These are also known and referred to as true positive and true negative rates.

When evaluating ML algorithms and model performance, we have to consider the erroneous conclusions as well as the correct classifications. If an algorithm is systematically repeating its imprecise conclusions, then this is

what is known as *bias*. Bias in humanities has come to carry a negative connotation, however in ML we can use bias for scientific evaluation. The main difference between humanitarian bias and scientific bias, however, can be found in the subjectivity and objectivity of the results. This is an important concept that needs to be assessed. In cases where statistical noise generates problems for the model, we call this *overfitting*.

Evaluation of model performance is the most important advance presented by ML and is carried out through the use of separate training and testing datasets. Models produced by ML algorithms are initially produced on training sets that are then tested on separate datasets in order to judge the observed performance. The best way to generate both types of datasets is through splitting our experimental samples to produce the training set, and the resulting testing set. In general a split ratio of 80% : 20% is employed (Brownlee, 2016a), however in cases where sample size is sparse and overfitting is likely, a more robust 70% : 30% split is preferred.

Alongside sensitivity and specificity values, the statistical efficiency of model performance can also be assessed through its kappa correlation values. The Kappa statistic adjusts accuracy when considering the possibility of a correct prediction by change alone (Lantz, 2013). A value of 0.8 in Kappa is considered as a powerful predictive model (Kuhn and Johnson, 2013).

Whereby considering the balanced accuracy of a model, the Kappa (> 0.8) and the sensitivity/specificity values (low number = high misclassification), the analyst can assess the efficiency and robusticity of each algorithms performance, choosing the best algorithm according to the test at hand.

ML algorithms in general can be divided into three main categories: supervised, unsupervised and reinforcement algorithmic models (Fig. 2.24). Quite simply, the supervision of a model implies the amount of the programmers input is involved into preparing the model. Supervised models usually consist in providing the algorithm with a well labelled data set, and through systematically mining the data, the algorithm is taught to find the most efficient means of arriving at the same conclusion. Supervised models are the most common type of ML algorithm and are usually employed in classification tests (Cortes and Vapnik, 1995; Kuhn and Johnson, 2013; Lantz, 2013; Yung-Wei, 2015). Unsupervised models are an advanced means of pattern recognition, whereby the system is provided an unlabelled dataset and searches for patterns without the input of the programmer. The most common means of carrying out unsupervised tasks is in the use of these algorithms for clustering (Kuhn and Johnson, 2013; Yung-Wei, 2015; Cholet and Allaire, 2017).

Reinforcement learning models are most often employed in DL, and can simply be described as a combination of the aforementioned learning techniques described above. Reinforcement learning is a complex application of AI where the algorithm is first left by itself to try and learn the data. The algorithm reacts with its environment (or task) deciding on an action to take, while observing how the environment changes with each decision. The reward signal produced, informs the algorithm how well or poorly it is performing. In some cases, an analyst is present to help inform the model when it is making a mistake or making good progress. Through this the algorithm adjusts its weights throughout the process of interacting with the task, the environment and the programmer all at the same time. These types of algorithms have been most successfully employed when teaching models specific tactics and real-time problem solving issues (Szepesvári, 2009; Nagabandi et al., 2017; Salimans et al., 2017; Gauci et al., 2019), such as how to play chess (Silver et al., 2017), and are especially successful in modern day robotics (Wang et al., 2012; Kalashnikov et al., 2018; Mahmood et al., 2018).

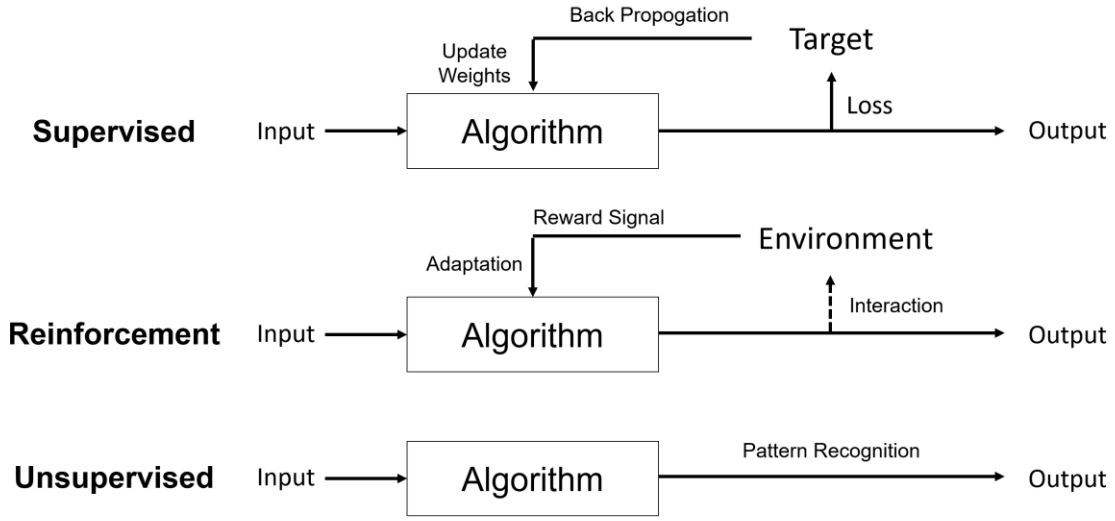


Fig. 2.24. – Figure summarising the different types of AI algorithms, including supervised, unsupervised and reinforcement learning. Figure by L.A. Courtenay, yet inspired by Wang et al. (2012)

The main difference between different learning models is found in the way the models update their weights. Algorithms are *stochastic* in nature, which means that they rely on randomness to initialise their weights before training. The use of randomness may sound contradictive, however it is highly useful for ensuring that the model is well adaptive to different datasets and is thus more robust when confronted with new unknown data (Brownlee, 2016a, b, 2019a; Chollet, 2017). The algorithm then proceeds to find the best way of computing links between the variables, using different self-evaluation techniques to adjust and correct each computation to provide the best results. The most popular techniques, especially in supervised learning, employ *backpropagation* and *gradient descent*. Both use notions of calculus and algebra to monitor the amount of errors produced for each time the algorithm is run. The next step that the algorithm performs is to then correct itself and reduce these errors. *Backpropagation* simply refers to the way data is fed back into the algorithm for weight optimisation, while *gradient descent* is a specific concept that tries to find the best parameters which produce the lowest errors.

The most famous and popular ML models to be developed are found in computational systems that replicate the architecture of the human brain in the form of artificial neural networks (Bishop, 1995; Günther and Fritsch, 2010). Other frequently employed and equally complex systems search for "hyperplanes" in non-linear data, mining for the most efficient means of discriminating between groups that would normally increase marginal errors in standard linear discriminant functions (Sing et al., 2005).

In recent years the field of ML has divided itself once again to include the sub discipline of DL models. DL was originally conceptualised in the 1980's (Dechter, 1986), yet wasn't truly developed until the early 2010s (Chollet and Allaire, 2017), and has proved highly successful in providing key breakthroughs that standard ML techniques could not have hoped to accomplish in the past. Among these include near human level classification of images and even speech recognition. DL generally requires long tuning processes and architectural remodelling according to data sets, yet achieve higher degree of classification and pattern recognition because of this. These systems essentially take standard ML techniques to an extreme, and can

be defined as a means of processing successive layers of representations (Chollet and Allaire, 2017; Patterson and Gibson, 2017), searching for the optimum mean of understanding the data. The depth of the model refers directly to the number of layers used in its construction and is most commonly found in the form of neural networks. A standard neural network tends to employ one hidden layer of neurons, whereas DL neural networks can use multiple. At times this may imply a longer training time and can even require complex mathematical calculations for its construction (Gu et al., 2018). The advantage of DL models, however, is seen when exposing them to the training data. DL models learn through their mistakes to fine auto-tune their parameters, ensuring that when exposed to real unseen data, their means of classification are much higher than *shallow* ML models (Chollet and Allaire, 2017).

In science, a combination of both ML and DL have proven to be one of the key advances of the 21st century. Hybrid statistical approaches have been used for in depth cancer research projects (Bellazzi and Zupan, 2008; Kim et al., 2012; Koopman et al., 2018; Sherafatian, 2018), providing new means of detecting and providing further hope for patients (Heredia et al., 2015; Tapak et al., 2018). Neural networks and supervised learning classifiers have provided a revolution for cancer survival predictions (Fernandes et al., 2018), as well as increased accuracy in diagnosis (Webb and Agar, 1992; Mangasarian et al., 1995; Richter and Khoshgoftaar, 2018). Pharmacists on the other hand have employed multiple algorithms in the preparation and study of chemical formulations for new projects (Ekins, 2016; Gawehn et al., 2016; Korotcov et al., 2017; Mamoshina et al., 2018; Yang et al., 2018), as well as predicting the success and properties of products new to the market (Han et al., 2018). DL's success when processing images have increased disease detection in biopsies (Tosun et al., 2009), also providing interesting insights into fields of engineering for material analysis (Zhang et al., 2017; Basu et al., 2018; Ruberto et al., 2018).

While most remain oblivious to the impact AI has on their lives, the most popular features of social media and marketing schemes are also structured using complex ML and DL algorithms. These include Google (Abadi et al., 2016), Facebook (Gauci et al., 2019; Loizou et al., 2019; Zhang et al., 2019), Netflix (Gomez-Uribe and Hunt, 2015), and most apps for mobile phones (Abadi et al., 2016). Mainstream applications of AI additionally build from the concept of Open Access and Open Source publishing, including the possibility for general public participation in computer science projects (Deng et al., 2009; Su et al., 2012; Deng et al., 2014; Russakovsky et al., 2013, 2015; Hempel, 2018). This can be observed to the point where institutions such as Cancer Research UK provide for Open Access participation in leading research, employing ML and DL algorithms for cancer detection and treatments[2].

With the development of free software such as Python (www.python.org, Python Software Foundation, 2018) and R (www.rproject.org, Core-Team, 2018), as well as the Open Access policies enforced by their developers when publishing (Su et al., 2012), AI has become a widespread science that almost *anyone* can participate in (Simonite, 2018).

AI has provided a new means of processing data, with incredibly high accuracy rates that are easily obtainable (Michalski et al., 1983). The introduction of AI in archaeology, however, has unfortunately occurred much later.

Arriaza and Domínguez-Rodrigo (2016) present the first case of ML algorithms applied to taphonomic data, using a combination of typical supervised ML

[2] https://www.cancerresearchuk.org/funding-for-researchers/how-we-deliver-research/grand-challenge-award/artificial-intelligence

algorithms to process features of carnivore bone accumulations in the Olduvai Gorge (Bed I, Tanzania). Here the authors employed a mixture of Neural NETworks (NNET), Support Vector Machines (SVM) and ML tuned classification trees to process multivariate taphonomic data of both a qualitative and quantitative nature.

This initial introduction of more complex statistical algorithms can be considered an important step towards a more robust means of processing data. Additionally, the use of multiple ethnological bibliographical sources for expert-based method validation inspires future possibilities for different means of approaching the archaeological record via both qualitative and quantitative attributes of a site.

The initial introduction of this approach, however, took a while to be developed further. In 2018 similar uses of supervised algorithms were used to process multivariate qualitative data proposed by Domínguez-Rodrigo et al. (2009a), providing more powerful classification methods using ML techniques (Domínguez and Baquedano, 2018; Domínguez-Rodrigo, 2018). Nevertheless, as already argued by these authors (Domínguez-Rodrigo et al., 2017), these variables are subjective in nature and a study of this type requires a much more objective means of data collection. Nevertheless, unsupervised cluster-based ML algorithms have also been employed in a revision of data provided by Egeland et al. (2018), demonstrating the potential of unsupervised AI techniques may have on future archaeological studies. A much more objective means of approaching taphonomic data employs the use of ML supervised algorithms for the classification of bone fracture patterns, providing a powerful use of objectively obtained quantitative variables, combined with qualitative observations, for the classification of the fracturing agent involved in an assemblage (Moclán et al., 2019).

More recently, the first use of Computer Vision techniques via Convolutional Neural Network processing of pixels in photographs have been able to provide very high classification rates when objectively confronting photographs of cut and trampling marks (Byeon et al., 2019).

AI in Archaeology however is not exclusive to taphonomic studies. New studies into the study of genome and DNA sequences between Neanderthal and Denisovan species have employed DL with exceptional results (Mondal et al., 2018). These results provide a much more efficient means of processing large strings of data, such as those presented by DNAmt and DNAnu structural sequences (Maxmen, 2018; Mondal et al., 2018), searching objectively and systematically for possible mutations in string sequences. Furthermore, an objective means of approaching fragment reintegration studies within restauration has also employed DL using neural network based computer vision algorithms (Derech et al., 2018).

Needless to say, the current understanding of AI in the archaeological register is relatively limited and currently superficial. The use of ML in processing qualitative data is an interesting introduction of powerful tools that can be used to improve our means of analysing data, however, more objective means of obtaining the data are needed if we intend for a more rigorous taphonomic understanding of site formation processes.

Throughout this work, I intend to propose a new more objective means of processing taphonomic BSMs, employing quantitative means of data collection alongside powerful AI based techniques in processing said information. In order to observe the impact of these finds on archaeological data, a preliminary taphonomic sample from the archaeological site of FLK-West has been used for method validation.

AI has the capability of revolutionising archaeological science, yet more investigation into this field may provide a paradigm shift for the future of studies into early Hominin populations of the Lower Pleistocene in Africa.

Chapter 3

Olduvai

3.1. The Olduvai Gorge

3.1.1. Location and General Context of the Gorge

The exceptionality of the archaeological and palaeontological complex of the Olduvai Gorge, Tanzania, has earnt its title as one of the most important windows into the history of human evolution. The number of sites and the temporal resolution presented in this area is one of a kind, presenting a chronological window, rich with archaeological material, ranging from 2.1 Ma to 15 Ka. The real value of this enormous concentration of sites, however, is the data the gorge has provided for human evolution and the behavioural and cognitive attributes of some of the first hominin populations of eastern Africa.

The gorge is located in the north of Tanzania, found between the edge of the Serengueti and the Ngorongoro crater and high lands (Figure 3.1.). The local landscape of the gorge presents a relatively flat stretch of open areas, dotted with many acacia trees. While many of the open areas are product of Masai deforestation, this particular area of Central Africa is distinctly more open than many of the more forested areas to the west of the continent. The Masai are the only current evidence of human activity in this area, yet for vast stretches across the landscape. To the south the great mountain scape of Lemagrut and the edge of the Ngorongoro occupies the majority of the landscape, while to the north, the tip of the Naibor Soit hills are visible, with slabs of local quartzite erupting from the ground. The gorge in itself is a relatively steep drop with a maximum depth of 100 metres. The sides of the

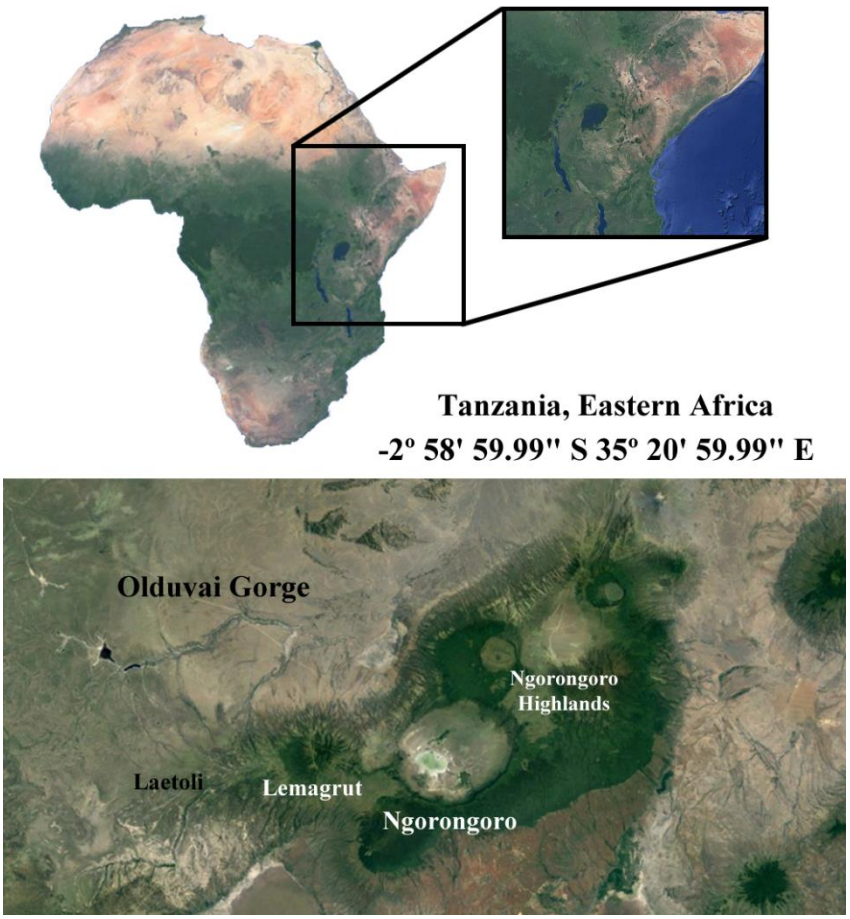


Figure 3.1. – Geographical location of the Olduvai Gorge, North Tanzania, Africa

gorge are covered in harsh vegetation, characterised mostly by sharp prickles and thorns and, most importantly, what the Maasai call *Oldupai* plants (*Sansevieria ehrenbergii*), giving the Gorge its name.

Providing a more general context, the Olduvai Gorge falls into the Great Rift Valley, a depression stretching from Mozambique to the Lebanon and measuring up to 6000 km in length, product of the separation between the Arabian and African tectonic plates. The East African Rift refers specifically to the area of this separation that includes Ethiopia, Kenya, Tanzania, Uganda, Rwanda, Burundi, Zambia, Malawi and Mozambique (Chorowicz, 2005). The development of this tectonic phenomena is currently dated to the Miocene, around 22-25 Ma (Ebinger, 2005).

The separation of these two plates has had a major impact on everything from the landscape to the vegetation of the affected areas. Much like any other tectonic process, the area contains a multitude of both active and dormant volcanoes, as well as long mountain ranges that separate a series of different ecosystems. From one side of the rift, the humidity brought in from the Indian Ocean produces stretches of dense vegetation, while to the other side of this natural barrier the landscape dries up giving way to the savannah. Such a change that can appear to be ecologically drastic in many occasions has a major impact on the wildlife and the ecological communities present.

The Savannah biome is an area of extreme trophic pressure in which species have a fundamental necessity to adapt in order to survive. This is especially relevant for primate populations, providing a context of particular interest for early Homo populations and their evolution.

.

3.1.2. Geology and Palaeoecology

The Olduvai Gorge is formed of 7 main beds, dating between 2.1 Ma to 15 Ka BP (Hay, 1976). While the initial characterisation of the Gorge's stratigraphy included only 5 beds (Reck, 1933), these were meticulously reanalysed and developed to include a total of 7 stratigraphic units, named Beds I to IV followed by Masek, Ndutu and Naisiusi (Table 3.1.). Bed I is the thickest and oldest layer, located at the very bottom of the Gorge, while the top three minor layers are much smaller in comparison and are named after the principal localities found in these beds. Bed I and II are characterised by similar depositional conditions represented by the presence of an alkaline salt lake. Around 1.3 Ma, the presence of the lake appears to vanish but is replaced by small ponds and marsh-like areas (Hay, 1976). Throughout the length of the stratigraphic sequence, multiple accounts of volcanic activity can be documented, creating layers known as Tuffs. These volcanic deposits are ideal for dating and have been able to provide a highly accurate absolute dating method that provides the best

Bed Number	Chronology (Ma)	Sites
I	2.1 - 1.7	FLK-Zinj, FLK-NN, PTK, DS, DK, MK
II	1.7 - 1.15	FLK-W, HWK, SHK, TK, BK
III	1.15 - 0.8	GRC, HG, HEG, JK
IV	0.8 - 0.6	HK, CK, CMK, Croc. K
Masek	0.6 - 0.4	FLK Masek
Ndutu	0.4 - 0.32	Ndutu Site, Nasera Soit, DGS
Naisiusiu	0.32 - 0.15	Naisiusiu Site

Table 3.1. - The Beds of the Olduvai Gorge. All dates are established using K-Ar radiometric dating (Hay, 1976). The sites and localities listed are considered referential sites for each bed, however many more sites could be included.

Tuff	Chronology (Ma)	Reference
IID	1.338 +/- 0.024	Domínguez-Rodrigo et al., 2013
IIC	Debatable	
IIB	Debatable	
IIA	1.66-1.74	Curtis and Hay, 1972; Manega, 1993
IF	1.803 +/- 0.002	Deino, 2012
IE	1.831 +/- 0.004	Deino, 2012
ID	1.839*	Deino, 2012
IC	1.848*	Deino, 2012
IB	1.848 +/- 0.003	Deino, 2012
IA	1.918*	Deino, 2012

Table 3.2. - Some of the dates for the main Tuffs of Beds I and II. The value of Tuffs IIC and IIB and their chronology are currently being debated (McHenry and Stanistreet, 2018). Dates marked with an asterix (*) are interpolated values (Deino, 2012).

means of chronologically understanding the sites presented in this archaeological complex (Curtis and Hay, 1972; Hay, 1976; Walter et al., 1991; Deino, 2012). These marker tuffs are numbered using the name of the bed in which they occur, followed by a letter indicating their position in relation to the other tuffs in the bed. Following this system, Tuff IIa is considered the oldest volcanic deposit in Bed II and Tuff IIb would be considered posterior (Hay, 1976). A summary of some of the main Tuffs and their dates for Beds I and II can be found in Table 3.2..

The general geology of this landscape implies the fluvial intervention that existed throughout Beds I and II to be of low energy, with the majority of sedimentation being product of clay decantation with frequent transformation by volcanic sediments (Hay, 1976). These clay deposits form up to 70% of the Gorge, suggesting a very low energy environment (Uribelarrea et al., 2014). Like most lake deposits in this area of Africa, the Olduvai palaeolake was formed by low energy waters flowing during seasons of heavy rain fall (Potts, 1996; Ashley et al., 2007, 2010). The abundant volcanic activity in this area produced a highly alkaline pH value within these waters (Uribelarea, 2018).

From a geological perspective, Beds I and II are practically identical, both consisting in a "monotonous" alternation of clay and volcanic tuff layers. The limit between the two beds is marked specifically by Tuff If, which is present and easily identifiable throughout the Gorge (Reck, 1933). Bed I is further characterised to include a Bed of lave flows, dated at 1.877 +/- 0.013 Ma (Deino, 2012), obtained via $^{40}Ar/^{39}Ar$ techniques (Deino et al., 2010).

The presence of tectonic faults are highly evident throughout the gorge, drastically changing the nature of the stratigraphic sequence. The gorge can be said to have 5 primary faults (Fig. 3.2.), with other minor faults located at specific points along the Gorge.

Beds III and IV are primarily characterised by a distinct change in the formation of the Gorge. Bed III for example is highly distinguishable due to its reddish-brown colour, and alongside Bed IV are considered to be deposits of an eolian nature. Above the Beds lie Holocene deposits of ash and alluvium.

The cutting of the gorge is believed to have been produced around 30 Ka, generated by a series of erosive fluvial events that eventually formed the Olduvai River (Hay, 1976). The gorge in general, however, is filled with posterior erosive events that have led to a distinct alteration of the stratigraphic sequence in some areas. This can be seen to such a degree that in some areas Bed Ndutu can be found to rest directly on top of Bed I.

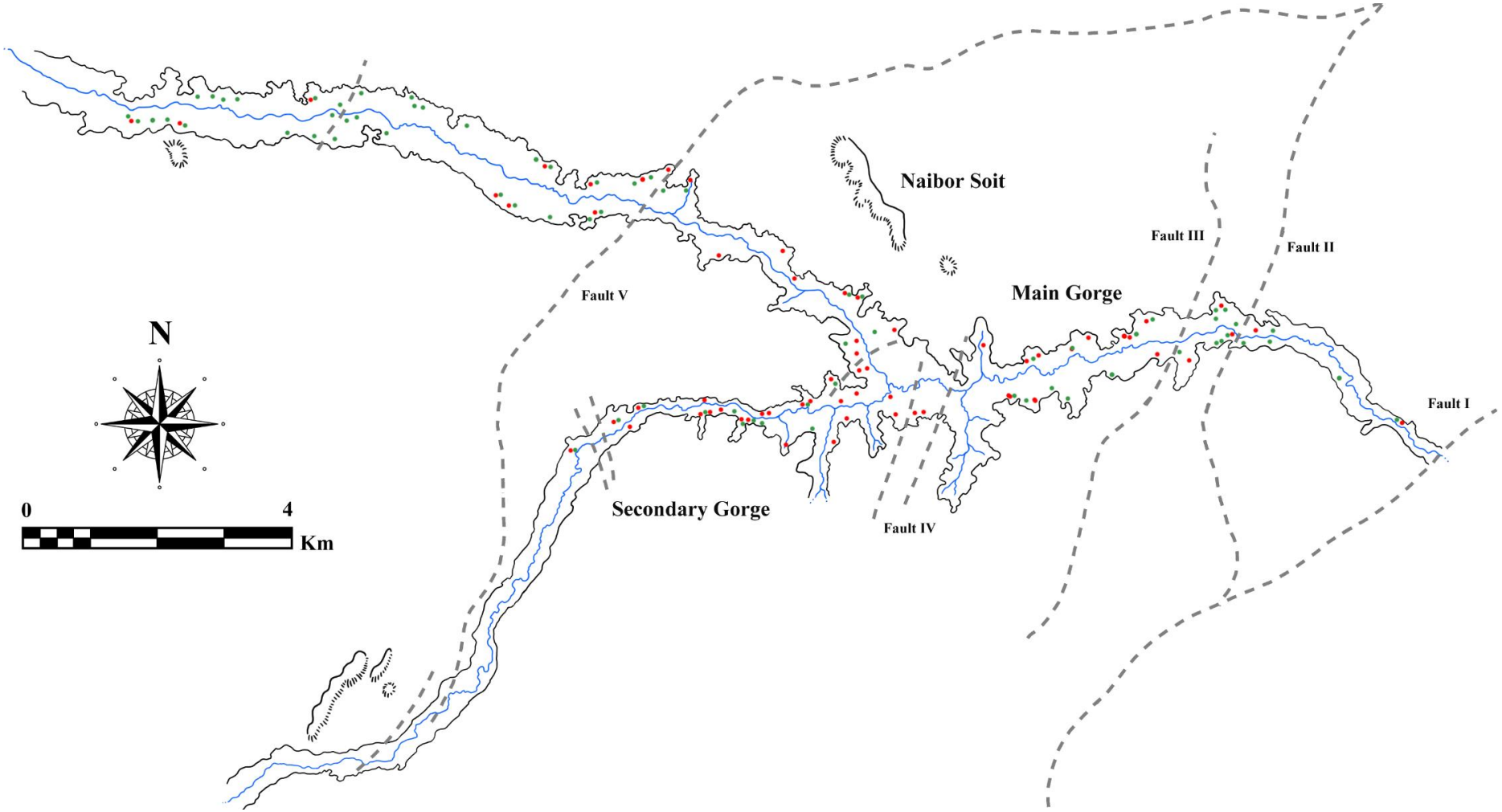


Figure 3.2. - Olduvai Gorge. Position of all the faults are derived from Hay (1976) and corrected using Uribelarrea et al. (2014). Green dots refer to Geological localities (Uribelarrea et al., 2014) and red dots refer to Archaeological localities (Leakey, 1965, 1971). Map and drawings done using a Wacom graphics tablet and in the ClipStudio software.

The Gorge in itself contains many localities of different natures, including archaeological, palaeontological and geological sites. The nomenclature of archaeological localities are designated using the initials of the person marking the first discovery (or in recent years in homage of someone) followed by the letters C, for cliff, K, for *Korongo* (the Swahili term for ravine), G, for gully and S, for site. According to this method, the site MLK stands for Mary Leakey Korongo. Further divisions exist including the geographical location of the site in reference to a main site. According to this method, the FLK (Frida Leakey Korongo) complex is divided into FLK-North, FLK-North-North (FLK-NN) and so forth. In some cases, where multiple sites of a different nature have been found, the sites are numbered, including the case of JK1, JK2 and JK3. Geological localities are numbered rather than lettered. Geological trenches located in the vicinity of another site are denoted by the letters a, b and c (etc.), including the example of 45, 45a, 45b, 45c, all located in the FLK complex. A glossary of the sites and their acronyms has been included in Appendix 1a.

The palaeoecology of this area has been topic of great discussion for quite some time, especially with regards to the Palaeolandscape of the FLK-Zinj layers where the *Zinjanthropus (Paranthropus) boisei* fossil [OH5] was found. Such a topic of interest is of great importance to evolutionary studies of early *Homo*, especially when trying to understand the development of our Genus and their coexistence with other bipedal primate species. From this perspective, interdisciplinary studies into the palaeoecology of these different landscapes have played a vital role in investigation of the sites present in Beds I and II.

Studies on the palaeolandscape associated with these populations highlight the significant importance of the palaeolake originally described by Reck (1933) and Hay (1976). These authors describe the lake to be of relatively large size in Bed I and to have reduced in size in Bed II, with an important density of archaeological sites orientating around the lake's shores (Leakey, 1971). The disappearance of the lake has been associated with the uppermost layers of Bed II, believed to be shortly after the formation of Tuff IId (Hay, 1976) and thus posterior to 1.338 +/- 0.024 Ma (Domínguez-Rodrigo et al., 2013a). The disappearance of the lake, however, does not directly imply a dry episode, yet is believed to indicate a landscape occupied mostly by marshlands as implied through the fauna present in the upper layers of Bed II (L. Leakey, 1965).

The ecology of the areas surrounding the lake through palaeobotanical evidence indicate Bed I to be characterised by the dominant presence of closed-wooded habitats populated mostly by woody plants, trees and shrubs (Ashley et al., 2010). Phytolith evidence has also been able to indicate an ecosystem abundant with palms (Piperno, 1988; Ashley et al., 2010) and supports the density of vegetation cover (Barboni et al., 2007; Blumenschine et al., 2012). However, caution has to be taken with such an interpretation due to the preservation of phytoliths in these contexts (Uribelarrea et al., 2014). Nevertheless, most studies conclude that the FLK-Zinj palaeolandscape of Bed I presents a mixed landscape with abundant vegetation and areas of dense coverage (Uribelarrea et al., *ibid*).

Bed II on the other hand marks a distinct change, represented especially in the zooarchaeological record towards the superior levels of this Bed (L. Leakey, 1965; Potts, 1988). Bed II begins to present a more open landscape, with patches of dense woodland (Arráiz et al., 2017). A more open landscape plays a vital role in archaeological studies concerning the much denser concentration of ancient hominin activity (Domínguez et al., 2007).

It is important to highlight, however, that over the stretch of 0.58 Ma the palaeolandscape of Bed II changed drastically. Bed II can be divided into different ecological sections. The first of these changes are marked by the Lemuta Tuff (Hay, 1976), presenting a change to a period of dryer spells, thus effecting

the lake. This change can also be observed around 1.5 Ma (Hay and Kyser, 2001) and a final change can be observed around 1.3 Ma, as previously stated with the disappearance of the palaeolake.

The important presence of low energy fluvial systems, however, is especially important in Bed II, as observed in multiple sites throughout this stratigraphic unit (Diez-Martín et al., 2017; Organista et al., 2017; Uribelarrea and Domínguez-Rodrigo, 2017; Uribelarrea, 2018).

Through these different changes we can observe how Bed II began with a much more aquatic ecology than the superior levels, characterised by the gradual disappearance of the lake, a constant formation of low energy fluvial formations and a mixed landscape of dense vegetation and open areas.

3.1.3. Historiography and Key Finds

The Olduvai Gorge has been fruit of numerous finds throughout the 20th and 21st century, providing a high quantity of sites rich with both human and faunal remains associated with abundant lithic tools.

The discovery of fossil remains in the Gorge have been documented ever since the 1910's by a team of German investigators led by the Geologist Hans Reck (1933). The unfortunate turn of global events produced by the First World War prevented Reck's work to go further, however after British control took over Tanzania, expeditions led by the Kenyan/British palaeoanthropologist Lewis Leakey continued to develop Reck's work with extensive campaigns and investigations throughout the entirety of the Gorge. To date the work of Lewis and Mary Leakey can be considered the most referential archaeological studies of this area (L. Leakey, 1965: Leakey, 1971; Leakey and Roe, 1994). Mary Leakey can quite possibly being considered one of the most important archaeologists in the field of Human Evolution.

Among the finds, the Gorge has been fundamental for the study of human evolution, producing a high quantity of hominin remains associated to almost 100 different individuals. Among these, fossil remains have been included as a holotype for numerous species, including *Homo habilis* (Leakey et al., 1964; Tobias, 1991), and *Paranthropus boisei* (Leakey, 1959; Tobias, 1967). Additionally, new finds of a > 1.84 Ma *Homo* sp. species is likely to indicate the possibility of a Hominin group that we are yet to discover in detail, coexisting with both *Homo habilis* and *Paranthropus boisei* populations in Bed I (Domínguez-Rodrigo et al., 2015b). Among these finds, the hominin species represented in the Gorge include *Homo sp.*, *Homo habilis*, *Paranthropus boisei*, *Homo ergaster/erectus* and even *Homo sapien* remains (Leakey, 1959, 1960; Napier, 1960; Leakey and Leakey, 1964; Leakey et al., 1964; Tobias, 1967; Day, 1971; Johanson, 1987; Almécija et al., 2010; Domínguez-Rodrigo et al., 2013a, 2015b). Another important site found near the Gorge includes the fossil footprints of Laetoli, considered the oldest direct evidence of bipedal features in hominid populations (Hay and Leakey, 1982; Leakey and Harris, 1987).

One of the most important features of these hominid fossils, however, consists in their coexistence across the same landscape, to the point where some sites present direct evidence of multiple hominin species (Leakey, 1971). Recent advances in genetic network analysis of different hominid fossils have revealed biological evidence to argue sexual interactions between *Paranthropus* and *Homo* species (Underdown et al., 2017), highlighting the likely context in which these multiple hominid groups would have interacted.

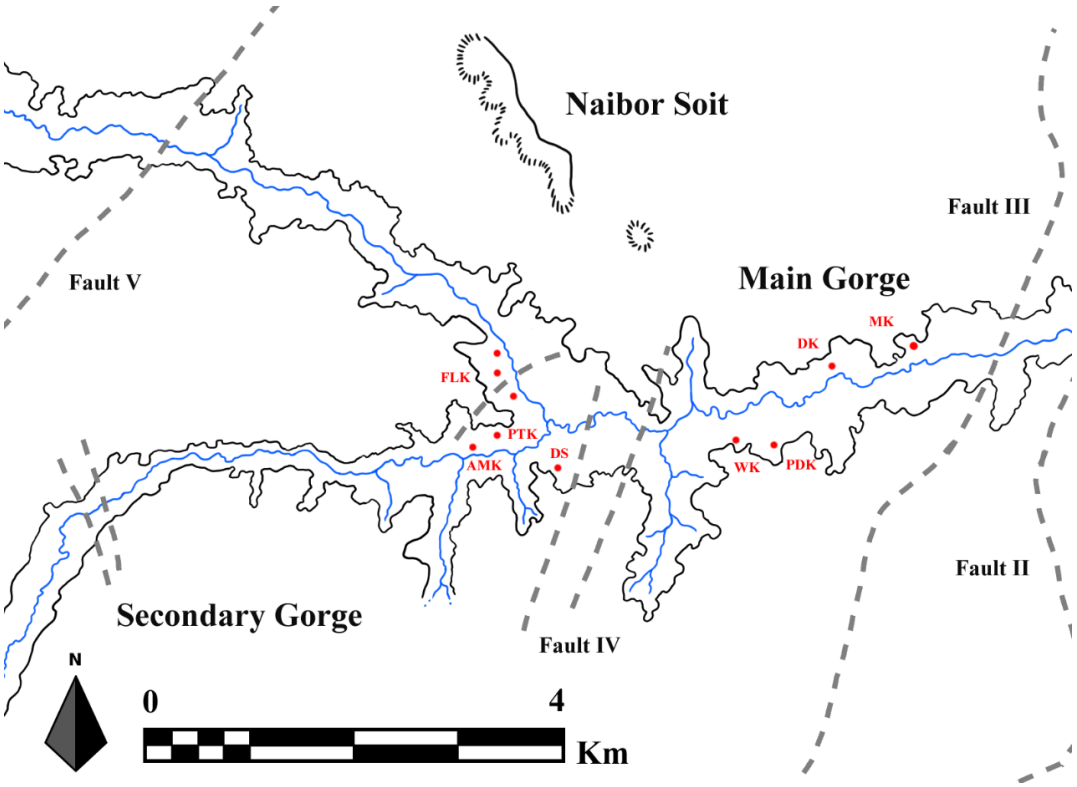


Figure 3.3. - Key Sites of Olduvai Gorge's Bed I. Map and drawings done using a Wacom graphics tablet and in the ClipStudio software

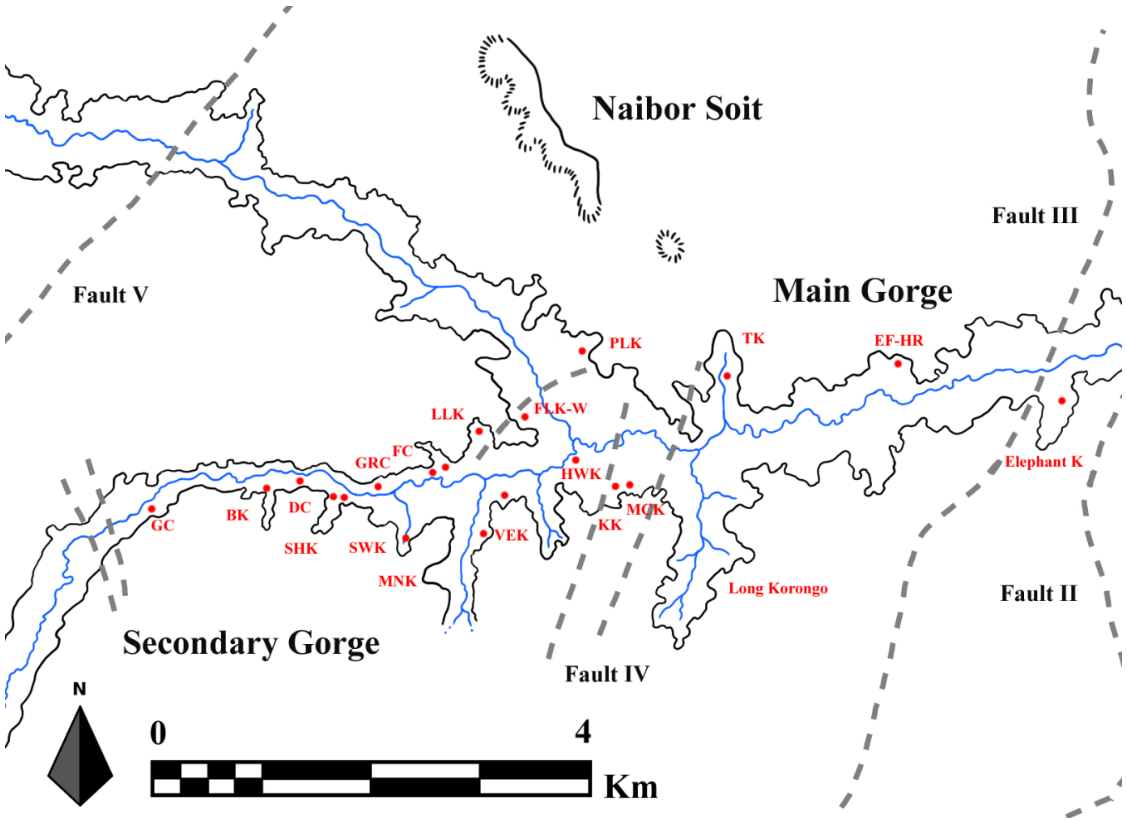


Figure 3.4. - Key Sites of Olduvai Gorge's Bed II. Map and drawings done using a Wacom graphics tablet and in the ClipStudio software

One of the most important features of the Gorge however is not solely found in the hominin remains present, yet the objects and finds found associated with them. The Olduvai Gorge provided the necessary lithic material for the characterisation of the Oldowan technocomplex (Leakey, 1971), as well as for sufficient material to resolve some key issues in the Hunter-Scavenger debate (Blumenschine, 1986, 1995; Domínguez-Rodrigo, 1997, 1999; Domínguez-Rodrigo and Barba, 2006; Blumenschine, 2007; Domínguez-Rodrigo and Barba, 2007; Domínguez-Rodrigo et al., 2007).

If we are to further consider the stratigraphic potential of this complex, stretching over almost 2 Ma, the numerous sites of Beds I through to the late Pleistocene deposits provide sufficient lithic materials to identify multiple cultural periods (Leakey, 1967, 1971), as well as the transitions between them (Leakey, 1967; Semaw et al., 2009; Diez-Martín and Eren, 2012; De la Torre and Mora, 2014).

Bed I (see Fig. 3.3.), for example, can be classified primarily through what has come to be known as the FLK-Zinj palaeolandscape. The original finds of the *Zinjanthropus boisei* fossil, OH5 (Fig. 3.5.), later named part of the *Paranthropus* genus of robust australopithecines (Leakey, 1959; Tobias, 1967), were initially identified to be associated with heavy duty and light duty simplistic tools. These tools alongside others found in Bed I were used to define the Oldowan technocomplex (Leakey, 1967, 1971).

The unique association of these simplistic tools with the *Paranthropus* genus, however, led the investigators to wonder whether these robust bipedal individuals were the true authors of these tools.

Soon after this discovery, in 1960 the team discovered more human remains in a site close to FLK-Zinj. The site of FLK-NN revealed new human fossils, including the almost complete foot of OH8 (Leakey, 1960), as well as an incomplete crania, OH7, that eventually became the holotype for a new hominin species; *Homo habilis* (Leakey, 1964).

Figure 3.5. - OH5 Skull, *Paranthropus boisei*. Leakey's "Dear Boy" from the FLK-Zinj site. Source: Smithsonian Museum of Natural History exhibition webpage

This discovery became a key find that helped unravel new data about the authors of the Oldowan technocomplex. Additionally, the association of these tools with abundant cut mark bones was sufficient in sparking off a new debate regarding their interpretation (Leakey, 1971; Binford, 1981; Bunn and Kroll, 1986; Blumenschine. 1986, 1987, 1995; Domínguez-Rodrigo, 1996, 2007, and many others).

Bed II (see Fig. 3.4.), on the other hand, presents a higher concentration of sites with a mixed technological nature. Multiple accumulations present an unclear transition between Oldowan and Acheulean technocomplexes, leading Leakey to describe a process of development known as the Developed Oldowan (DO) type A and DO type B (Leakey, 1967). Some key examples of DO sites include BK (Diez-Martín et al., 2009; Sánchez-Yustos et al., 2016), and SHK (Diez-Martín et al., 2014; Sánchez-Yustos et al., 2016, 2017a), both associated with a very rich faunal assemblage (Domínguez-Rodrigo et al., 2009b, 2013b, 2014a; Yravedra et al., 2017b; Organista et al., 2016; Organista, 2017; Organista et al., 2017), and in some cases hominin remains as well (Domínguez-Rodrigo et al., 2013a). Other sites on the other hand present clear evidence of Large Cutting Tools (LCTs) including Bifacial Handaxes and

Cleavers, characteristic of the Acheulean technocomplex.

These mixed cultural movements in hominin populations at this point implicate an increased cognitive complexity present within these populations (Stout et al., 2000, 2008; Stout and Chaminade, 2007; Higuchi et al., 2009; Lepre et al., 2011; Lewis and Hamand, 2016; Putt et al., 2017; Petö et al., 2018; Toth and Schick, 2018), typically assumed to be associated with *Homo ergaster/erectus* populations present within the Gorge (Leakey, 1964; Diez-Martín, 2005; Maslin et al., 2015). The dynamic range of different handaxes ranging from Beds II to IV also present a clear change in cognitive capacities of the associated hominin populations, with a significant increase in quality of tools in Bed IV (Leakey, 1971).

Acheulean sites of Bed II present a number of highly important sites, including the immense lithic assemblage of TK (Leakey, 1971; Santonja et al., 2014; Rubio-Jara, 2017; Santonja et al., 2018), presenting a wide variety of different handaxes ranging from very large implements in the Lower Occupation floor, TKLF, to smaller dimensions more typical for cutting and butchery activities in the Sivatherium Floor, TKSF (Rubio-Jara et al., 2017; Santonja et al., 2018).

Finally, new excavations in the lower levels of Bed II near the FLK-Zinj site were able to reveal a highly interesting concentration of an Acheulean tecnocomplex associated with faunal remains. These finds will be explained in continuation and presents the case study for this Master's Thesis.

3.2. The Site of Frida Leakey Korongo West (FLK-West)

The Frida Leakey Korongo is a ravine located on the left hand side of the main Gorge, close to the junction where the two branches of the Olduvai Gorge meet. The Korongo in question is named after the woman who discovered it, Lewis Leakey's first wife; Frida. This Korongo in particular holds some of the most important finds in palaeoanthropological history and has provided a great wealth of data regarding hominin origins and early behavioural traits.

The FLK complex (Fig. 3.6.) was initially discovered in 1931 and consists of a series of sites, most of which are located in the bottom of the ravine, thus forming part of Bed I. While the majority of remains are located here and have thus occupied the majority of investigation throughout the years, in Mary Leakey's first description of this complex she did indeed mention the presence of stone implements and fossils in several layers at Bed II. In this initial description and mention of possible Bed II layers in this area, M. Leakey continued to describe "*a trial excavation in the upper part of Bed II, carried out in 1960-1, yielded an assemblage of Chellean artefacts*" (M. Leakey, 1965: 103).

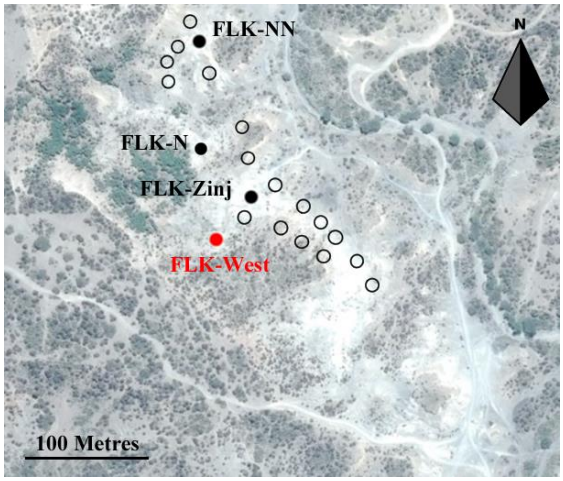


Figure 3.6. - The Frida Leakey Korongo. Leakey's original sites are marked in black, while the sites excavated by the TOPPP team are marked as black circles (Derived from Uribelarrea et al., 2014). FLK-West is marked in Red.

In the March of 2012, a team of investigators began a general revision of the original Leakey sample in the Dar es Salaam National Tanzanian Museum. Up until this date, the oldest Acheulean accumulation to be known

Standard Age and Decay Constant	**FLKWa**	**FLKWb**
Steiger and Jager (1997) Renne et al. (1998)	1.681 +/- 0.015	1.648 +/- 0.019
Renne et al. (2011)	1.698 +/- 0.015	1.664 +/- 0.019

Table 3.3. - Dates in Ma of Tuffs FLKWa and FLKWb using the 40Ar/39Ar dating technique. Decay constants used in the dating process are cited to the left. Data from Diez-Martín et al. (2015) Supplementary Appendix 1.

within the Olduvai Gorge was that of EF-HR, situated in the Middle Bed II, estimated to date around 1.6 Ma (Leakey, 1971). Inspired by the collections in the museum, and through following the notes and finds of Mary Leakey's original excavations in the 1960's, members The Olduvai Palaeoanthropological and Palaeoecological Project (TOPPP) team organised a series of prospections in the area, with the intentions of searching for the origin of the Acheulean. In June, 2012, the TOPPP team discovered a site, a little to the west of the original FLK-Zinj, rich in Acheulean material (Diéz-Martín, 2018).

The new site, appropriately named FLK-West, has been sporadically excavated between the years of 2012 and 2015, with the field season of 2018 spent expanding the area of the site and preparing future excavation surfaces to begin work in the September of 2019.

The stratigraphy of this site has been able to identify up to 6 layers (Fig. 3.7.), located in between two tuffs positioned about a meter above Tuff If, thus establishing this site to be a member of Bed II. These tuffs have been named FLK-Wa and FLK-Wb, both being dated via the $^{40}Ar/^{39}Ar$ radiometric absolute dating technique in the Scottish Universities Environmental Center (Glasgow, Scotland). Both Tuffs situate the FLK-West assemblage at an approximate age of just under 1.7 Ma (Table 3.3.), classifying this assemblage as one of the oldest Acheulean accumulations to date (Diéz-Martín et al., 2015).

Geologically speaking, the site is situated in a fluvial palaeochannel that would have run North-South from the Lemagrut slopes to the Olduvai palaeolake. Recent studies into these canals have found that the section corresponding with FLK-West consists of a total width of 40m and maximum depth of 1.2m. The edge of the canal is much shallower, with less than 50cm in depth. Each of the six layers are interpreted as different flood events, depositing sands of different sizes and granular composition (Diez-Matín et al., 2015).

The formation of a site in such a context is hard to imagine, especially considering the open air nature of the Gorge. The remains observed in FLK-West of both lithic and osteological nature do not present a preferential orientation, while lithic and bone refits are also evident throughout the site. Such an ideal spatial distribution is highly indicatory that the fluvial processes produced by the palaeochannel were of low energy, thus providing a higher preservation rate for the site in general (Uribelarrea, 2018). Additionally, in order for the water to not have moved the majority of the remains, the velocity and force of the current has to have been strong enough to move large amounts of sediment for a quick deposition over the remains, yet not a strong enough current to move the remains themselves (Uribelarrea et al., 2017).

Extensive investigation across the gorge into the geomorphology of the landscape have been able to characterise the nature of these canals over a surface area of 1.5 km (Uribelarrea et al., 2017). Analysis of this landscape takes into consideration what is known as the Lower Augitic Sandstone (LAS) depositional environment, originally described by Hay (1976: 60) as an augite-rich sandstone deposit which discomformably overlies the Lemuta Member between FLK and the fault at KK. In these studies, analysts have been able to

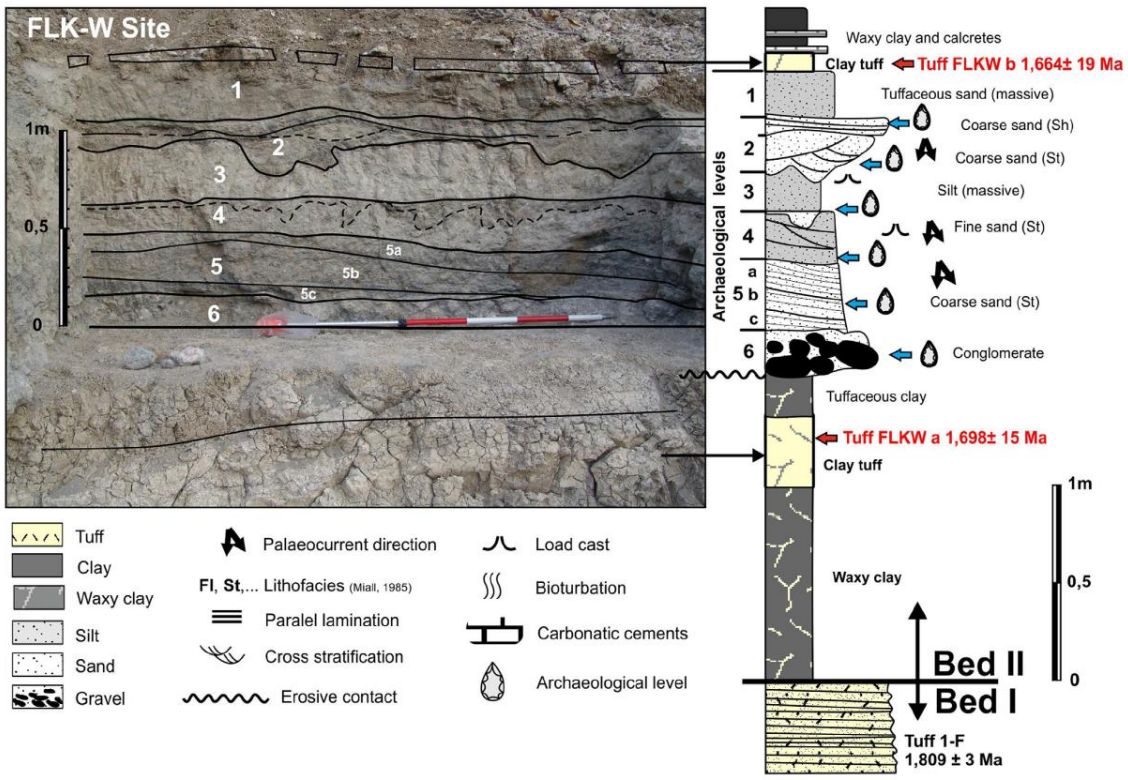


Figure 3.7. - Stratigraphy of FLK-West. Source: Diez-Martin et al. (2015)

detect three different types of palaeochannels in the area (Fig. 3.8.); the first consists in an alluvial fan, the second in a series of rivulets creating a criss-cross pattern and the last in multiple sinuous channels (Uribelarrea et al., 2017). Considering stratigraphic data that confirms a contemporary position with Oldowan and DO sites at HWK (HWK, HWK-East and HWK-EE) confirm that FLK-West presented a denser cover of vegetation located closer to the Olduvai palaeolake shoreline than

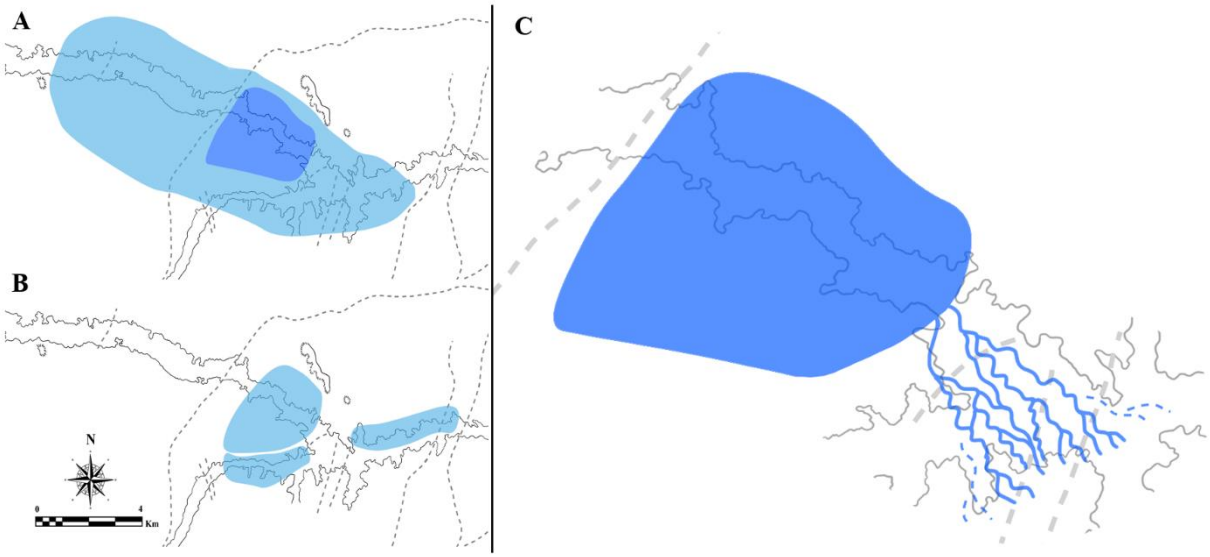


Figure 3.8. - Description of the Paleolakes and Fluvial Environment of the Lower Beds of the Olduvai Gorge. (A) Bed I. Light blue presents the flood plain limits of the palaeolake. (B) Upper Bed II. A Series of different temporal pools and marshlands. (C) The paleochannels present between HWK and FLK-West, draining in to the paleolake that is still present in lower Bed II. Data for A and B from Hay (1976). Data for C from Uribelarrea et al. (2017). Drawings done using a Wacom graphics tablet and in the ClipStudio software

sites located further to the South and South East of the Gorge.

The Acheulean lithic material from FLK-West (n = 2120) are highly concentrated in the lower levels of the site. Most are made from Naibor Soit Quartzite (73.67%), while the the other materials present in FLK-West, consist in volcanic rocks (18.2%) and chert (7.0%).

The Quartzite found in the multiple sites across all levels of the Gorge can be sourced mostly from the Naibor Soit hills, where this material appears in great slabs protruding from the ground. The Gorge, however, is also abundant with volcanic rocks, including basalt and phonolite, and also cherts (especially from a source found in the site of MNK). Chert nodules however appear in small quantities and irregularly shape, not ideal for knapping. Quartzite on the other hand is ideal considering its abundant quantity.

Issues with the geological nature of the Naibor Soit quartzite have existed for quite some time with most authors characterising this material as a quartz when considering the texture and granular nature of this material (Blumenschine et al., 2008; de la Torre and Mora, 2013; de la Torre et al., 2013). While some class this material as a quartzite considering its metamorphic origin (Santonja et al., 2014). For the purpose of this study, the term quartzite will be used.

Morphofunctional studies of the lithic assemblage at FLK-West identify multiple layers of different technological groups across the site (Sánchez-Yustos et al., 2017b, 2018). In Layers 1 and 3 percussion implements are the most representative while levels 2, 4 and 5 present evidence of nuclei and small flakes. Level 6 however presents a great quantity of LCTs including bifacial handaxes, typical of Acheulean technocomplexes (Sánchez-Yustos et al. 2017b). While there is a clear separation between the lower and upper layers of the FLK-West lithic assemblage, authors believe that this is a product of the bias represented through the *Chaîne Opératoire*, still being attributed to the same cultural group of hominins (Sánchez-Yustos et al., 2018).

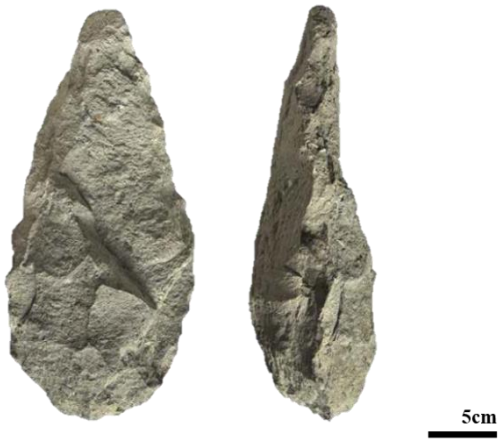


Figure 3.9. - The Great Basalt Bifacical Handaxe of FLK-West. Source (Diez-Martín et al., 2018)

Nevertheless, the great basalt bifacial handaxe of FLK-West (Fig. 3.9.), found in level 6, can be considered a great anomaly in comparison with the rest of LCT materials from this level. Measuring a total of 300.2 x 139.9 x 93.07 mm (Length x Width x Thickness) and a weight of 3.66 kg, this piece stands out amongst the other LCTs of FLK-West (179 x 96 x 64 mm; 1.3 kg). The biface is believed to have been knapped in a total of 12 different series of 3 knapping sequences. The first consists of 3 different sessions where invasive extractions and centripetal rotation of the piece allowed for volume reduction of the main face of the handaxe. The second sequence through another 4 sessions carried out on the other face of the tool intended to configure the final morphology. Finally, the last 5 phases of the third sequence consisted in the final touches with much smaller removals and less invasive extractions (Diez-Martín et al., 2018). The general shape of the handaxe consists in a clear almond shape with significant bilateral symmetry.

Considering the preparation and nature of the final symmetrical product, it can be argued that the great basalt handaxe of FLK-West presents a clear representation of a complex preparation process that is not accidental in nature, highlighting an advanced cognitive complexity which would fit in with

concepts of Acheulean population traditions. If we were to further consider the age of this piece, and the advanced nature of other LCTs in this assemblage, it could be argued that the FLK-West accumulation in general represents a group of hominins with high cognitive abilities.

The faunal assemblage associated with these LCTs constitutes a total of NISP of 1043. 138 of these remains are identifiable taxonomically while the remaining 905 can be grouped according to size. A total MNI of 65 has been calculated across all 6 levels (Table 3.4.). 61 of these individuals are herbivores, 2 are amphibious and 1 is part of the carnivore family.

Across all 6 levels, the number of Bovid individuals is the most dominant of all the taxonomic families with a predominance of Alcelaphines. Suidae species are also represented across most levels, followed by Equidae. To a lesser extent we find size 5 animals such as the Rhino, Hippo and Pelorovis, finally including proboscideans in the uppermost layers of FLK-West (Table 3.5.). Nevertheless, considering the number of indeterminable finds (NR = 612), a large percentage of the accumulations in each level indicate a well-represented number of Size 3 individuals, predominantly of Size 3a (Table 3.6.).

In general, dentition is the most represented element of the entire assemblage, however, considering the preliminary sample's small size, not too much can be said about these values (Yravedra and Domínguez-Rodrigo, 2018).

Level I can be characterised ecologically through macrofaunal remains in its presence of amphibian species and animals typical of aquatic habitats, combined with a high presence of open habitat species. Level II and III on the other hand are the least rich assemblages of all the site with little that can be said about the type of ecosystem represented. Level IV is characterised especially by animals of an open habitat, similarly presenting a high presence of animals accustomed to spending large amounts of time near or in the water. Level V is similar with a high quantity of open space animals, however, a particular change in this level presents a Tragelaphine which is more indicatory of wooded areas as well. Nevertheless, this level is predominantly characterised by open area animals. Finally Level VI also presents a high number of open area species.

Preliminary taphonomic studies at this current data have been able to highlight a high number of water abraded surfaces, product of the lower energy fluvial contexts presented by this site. Nevertheless, cut mark cross sections were able to reveal data indicating an important use of quartzite tools as cutting implements (Yravedra et al., 2017). Some preliminary observations by Courtenay (2017) found processing these marks difficult, product of their superficial nature, and topic of study for this particular Master's Thesis.

It is important to highlight, however, that to date a small number of excavation campaigns have been carried out on this area, producing relatively insignificant samples for an interpretation on the nature of the first Acheulean populations. Nevertheless, this Master's Thesis intends to use the small taphonomic sample that is currently available as a validation method in testing out the methodological approaches presented here. In continuation the taphonomic sample selected will be explained with the methods presented to study them. Additionally, an experimental sample is presented that provides a valuable analogy that can be used to decipher these taphonomic traces.

	Level I		Level II		Level III		Level IV		Level V		Level VI	
	NR	MNI	NR	MNI	NR	MNI	NR	MNI	NR	MNI	NR	MNI
Avifauna	1	1										
Pelusios sp.									1	1		
Crocodylus sp.	1	1										
Connochaetes sp.									4	2		
Megalotragus sp.			2	1			1	1	1	1	3	1
Alcelaphini Size 3a	2	2			1	1	1	1	5	2	6	1
Alcelaphini Size 3b									1	1		
Alcelaphini Size 2	1	1					5	2	3	1		
Hippotragus giga									1	1		
Tragelaphini sp.									1	1		
Kobus sp.							1	1				
Antidorcas marsupialis							1	1				
Gazella sp. Size 1							4	1				
Gazella sp. Size 2	1	1					4	1	1	1		
Syncerus sp.	1	1									1	1
Pelorovis oldowayensis									1	1		
Giraffidae									1	1	1	1
Kolpochoerus sp. Size 2									3	1	1	1
Kolpochoerus sp. Size 3	1	1					1	1	2	1	3	1
Metridiochoerus sp.									2	1	6	2
Suidae Size 3			1	1			1	1	1	1	1	1
Elephas sp.	2	1	2	1	2	1	1	1				
Equus oldowayensis	2	1							9	2	8	1
Hipparion cornelianum	1	1									9	3
Rhinocerontidae									1	1		
Hippopotamus sp.	7	2	1	1			10	1				
Crocuta crocuta									1	1		
TOTAL	**20**	**13**	**6**	**4**	**3**	**2**	**30**	**12**	**39**	**21**	**39**	**13**

Table 3.4. – MNI and NR of Different Species from the Faunal Assemblage of FLK-West. Data derived from Yravedra and Domínguez-Rodrigo (2018)

	Level I		Level II		Level III		Level IV		Level V		Level VI		TOTAL	
	NR	MNI	NR	MNI	NR	MNI	NR	MNI	NR	MNI	NR	MNI	**NR**	**MNI**
Bovidae	5	5	2	1	1	1	17	8	18	11	10	3	**53**	**29**
Girafidae									1	1	1	1	**2**	**2**
Suidae	1	1	1	1			2	2	8	4	11	5	**23**	**13**
Equidae	3	2							9	2	17	4	**29**	**8**
Rhinocerontidae									1	1			**1**	**1**
Hippopotamidae	7	2	1	1			10	1					**18**	**4**
Elephantidae	2	1	2	1	2	1	1	1					**7**	**4**

Table 3.5. – NR and MNI of faunal remains from FLK-West, grouped according to taxonomic family. Data derived from Yravedra and Domínuez-Rodrigo (2018)

	Level I	Level II	Level III	Level IV	Level V	Level VI
Size 1				5	4	5
Size 2	2		3	20	45	24
Size 3	8	4	6	22	53	35
Size 3a	3		6	11	39	30
Size 3b	2		1	7	22	46
Size 4	12	2	1	17	40	36
Size 5	12	1	2	11	7	12
Size 6			1	2	3	2
Indet.	17	18	14	84	105	103
TOTAL	**56**	**25**	**34**	**179**	**318**	**293**

Table 3.6. - NR of unidentifiable remains arranged according to African Size Group. Data derived from Yravedra and Domínguez-Rodrigo (2018)

Chapter 4

Sample and Methods

4.1. Experimental Samples

Considering how the purpose of this study is to provide new methodological framework for the study of BSMs, multiple experimental samples were used for statistical testing. The additional objective of providing a powerful analogical comparative sample with the FLK-West materials also required a large series of experimental marks for an efficient comparison and classification. To meet these objectives, a wide range of different BSMs were studied, using experimental materials as well as revising already published materials in order to build such a frame of reference. In continuation, all the samples produced, revised and studied will be explained in detail. Additional technical details have also been provided regarding the means of processing BSMs. Towards the end of this section, the statistics employed for both GMM, ML and DL have also been described in great detail.

4.1.1. Trampling Marks

A total sample size of 281 trampling marks were selected and processed. These marks were procured from two separate highly controlled experimental samples, the first sample comes from Yravedra (2005)'s PhD Thesis and the second from Domínguez-Rodrigo et al. (2009a)'s studies into differentiation between cut marks and trampling marks.

Yravedra (2005)'s trampling sample was produced under experimental conditions, using sieved sediments from levels 2 and 3 from the archaeological site of Peña de Estebanvela (Segovia, Spain). Sediment samples obtained from this site contain a mixture of compact and loose quartz sandy sediments. The composition of these sands consists of angular quartz granules with an average size of 250 µm. These sands were used to cover 14 axial and appendicular bones of an *Ovis aries* skeleton. Of the bones used, marks were produced on mandibulae (Number of Elements (NE) = 2), femorae (NE = 2), tibiae (NE = 2) and radiae (NE = 2). All bones were meatless when buried and both the bones and sediments were in a dry condition. These bones were then exposed to 5 minutes of trampling by a single, average weighted individual before being uncovered, cleaned and studied.

The total number of marks produced from this initial experiment consisted of up to 56 trampling marks. 30 marks present clear morphologies that could be processed using Courtenay et al. (2017)'s GMM model. This small sample was sufficient in preparing the methodological approach that would be consequently used when comparing trampling marks with cut mark samples.

Domínguez-Rodrigo et al. (2009a)'s sample is of a much larger size, and was carried out under multiple different experimental conditions. These experiments were carried out using cervid bones obtained from a legal organised hunting parties. These bones consist of a mixture of axial and appendicular skeletal elements, including femora, tibiae, radii, ulnae, humerii, vertebrae, ribs and scapulae. Before trampling experiments could begin, the meat was removed using metal knives, and then sectioned into smaller pieces using an electric saw. Each consequent piece was then examined and inspected for BSMs before the trampling experiments could begin. The electric saw enabled the bones to be sampled without creating additional striae that would have been evident if using percussion to break the bones. The consequent number of fragments included 47 axial elements and 129 appendicular elements.

The next variables controlled in this sample included the type of sediment used. Sediment samples consisted in fine-grained

sands (60 – 200 µm), medium-grained sands (200 – 600 µm), coarse grained sands (0.6 – 2 mm), a combination of different sand types in a clay stratum and finally gravels (> 2 mm). The second variable controlled was the time exposed to trampling. Two different experimental sets were carried out, the first with brief exposure to trampling (10 s) and the second with prolonged trampling (2 min).

An additional variable controlled included the individual producing the trampling. All individuals were young and of varying weights. All wore the same footwear so as not to create statistical noise in the sample.

The final variable controlled included whether the bones were buried when fresh or dry.

After the different trampling experiments, these bones were finally cleaned using a boiling neutral solution of water and soap. The consequent sample of trampling marks included 251 clearly identifiable marks with full present morphology that could be processed using Courtenay et al. (2017)'s GMM approach. This larger sample was sufficient in observing the nature of trampling marks under different sedimentary contexts and formative conditions.

4.1.2. Cut Marks

A wide range of experimental cut mark samples are currently available, considering raw material (Maté-González et al., 2016; Courtenay et al., 2017, under review; Yravedra et al., 2017a, b; Moclán et al., 2018), tool type (Courtenay et al., 2017) and cutting angle (Courtenay et al., 2018a). For this study, only cut marks produced by raw materials and tool types common from the FLK-West site were selected. These included quartzite from the Naibor Soit, local quartzite from the Gorge itself, and basalt from the same region. The tool types employed were mostly simple flakes, however retouched flakes and bifacial handaxes were also used. All of these samples come directly from Courtenay et al. (2017, Under Review).

For the comparison of cut marks with trampling marks, solely simple flakes were used. Tool type and raw material were included afterwards to classify the precise nature of the cut marks, once they had been classified as anthropic in nature.

The cut mark samples available from the aforementioned sample consists in 314 cut marks produced by different quartzite samples of different granular composition and size. These range from an average size of 2.64 x 2.87 mm to 5.00 x 5.00 mm, all originating from the Naibor Soit and surrounding areas of the Olduvai Gorge. These have been studied, measured and characterised using mineralogical techniques with thin sections in cross-polarised light (an example of this is provided in Fig. 4.1.) The eventual quartzite selected was that which appears most commonly in Bed II sites, especially in FLK-West. This sample presents a tight interlocking network of large quartz crystals, mostly accompanied with orientated prismatic muscovite crystals (Courtenay et al., Under Review).

As can be seen in Fig 4.2. and 4.3., these quartzite flakes present a highly irregular cutting edge, formed from large angular crystals. The average angle of the cutting edges used are approximately 60º (Fig. 4.4.). A total of 10 simple flakes were produced for this experiment by a professional lithics specialist, experienced in the *Chaîne Operatoire* of Acheulean sites.

Cut marks were produced on 2 *bovid* and 6 *suid* long bones. These bones consisted of femora (*bovid* = 2, *suid* = 2), tibiae (*suid* = 2) and radii (*suid* = 2). Only the diaphysis of appendicular skeletal elements were selected, considering these are the most abundant from the FLK-West sample. Nevertheless, even if different anatomical elements from animals of different sizes were used, these variables have previously been proved to be insignificant in

GMM studies (Maté-González et al., 2019). Bovid bones were obtained from a Halal butchers in Tarragona while *suid* bones were obtained from a local butchers in Madrid. All incisions were performed across the longitudinal axis of the bone while scraps of meat were still present. All bones belonged to adult individuals. These bones were then cleaned in a mixture of neutral boiling water and soap.

A final preliminary experimental sample was performed to generate cut marks that could be compared using bidimensional methods with trampling marks. This experiment consisted in the butchery of an adult *cervid* carcass from the Boumort National Hunting Reserve (Lleida, Spain). These experimental marks were produced with flint flakes from the Morero region of Northern Spain. The nature of these marks are much thinner, product of the finer grained raw material. These were able to provide a more valuable analogy when comparing their biometrical characteristics with those of trampling mark experiments from the Olduvai Gorge. Through this, additional comparisons could be used between trampling marks and cut marks of a different nature. 30 marks were processed from the total sample of 48 cut marks, due to the number of landmarks that could be successfully localised.

In the case of all experiments, the cutting edge was carefully examined to ensure that no blunting was occurring, thus ensuring that the attrition of the tool was not a conditioning factor in mark morphology.

The final experimental sample to be compared with FLK-West consisted of a total of 80 cut marks. 30 cut marks were produced using simple flakes in this experiment, to be lumped with 50 cut marks produced using the same tool type and raw material from Courtenay et al. (2017)'s sample.

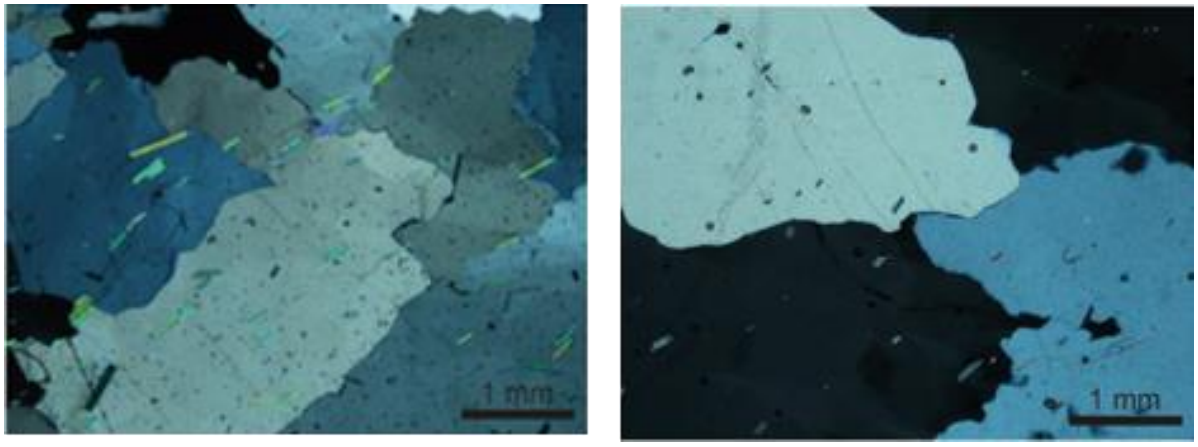


Figure 4.1. - Thin sections of two quartzite samples in cross-polarised light. Figure by David M. Martín-Perea and to be published in Courtenay et al. (Under Review).

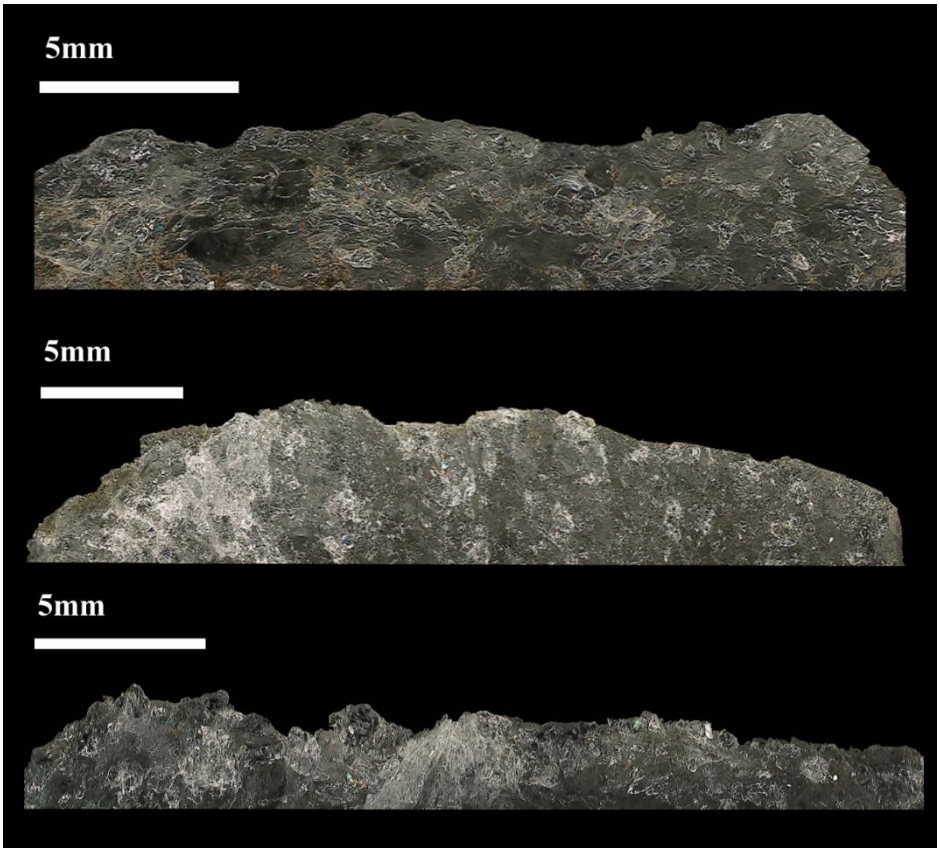


Figure 4.2. - Fine deatiled photos of three quartzite flakes used to produce experimental cut marks in this study. Photos were taken using the HIROX KH-8700 3D Digital Microscope using the Low Range lens at 35 x magnification

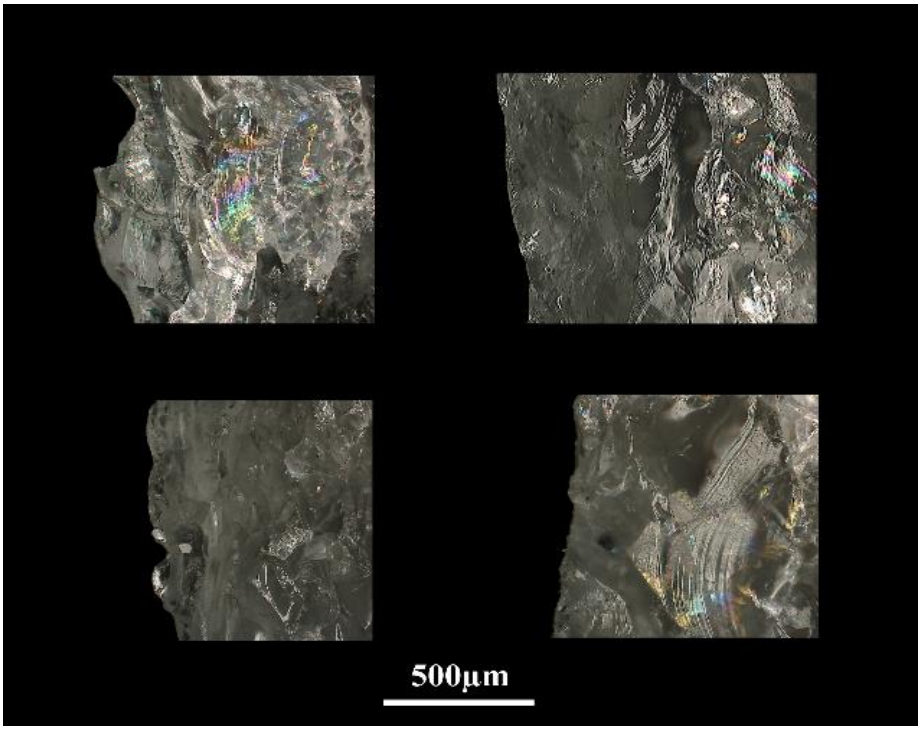


Figure 4.3. - Detailed photos of some quartzite simple flake edges, taken using the HIROX KH-8700 3D Digital Microscope. Photos taken using the Mid-Range lens at 200 x magnification

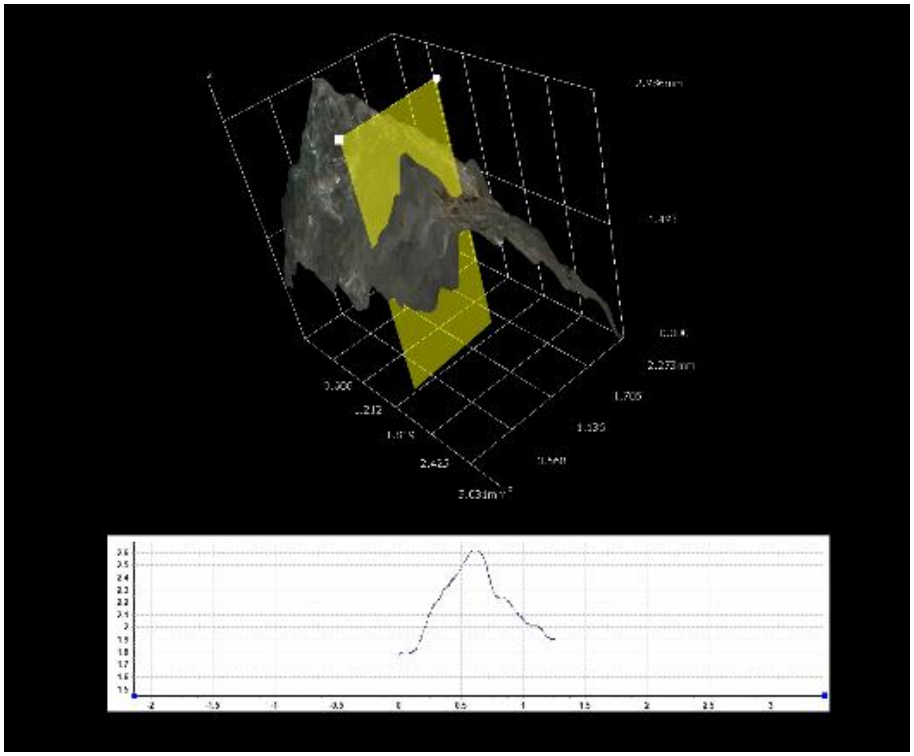


Figure 4.4. - Figure presenting a 3D model of a quartztie flake cutting edge, taken using the HIROX KH-8700 3D Digital Microscope.

4.1.3. Tooth Marks

Tooth marks included in this study consist of 89 tooth pits and 127 tooth scores, all produced by multiple (MNI = 2) adult carnivore individuals. The compared taxa included two African taxa including the spotted hyena, *Crocuta crocuta* (pits = 21, scores = 33), and the lion, *Panthera leo* (pits = 24, scores = 30). Due to current availability of African canid marks, the grey wolf, *Canis lupus* (pits = 24, scores = 30), was used for this study. Likewise, no sample of Leopard tooth marks are currently available for study, therefore this was replaced by another felid; the South American jaguar, *Panthera onca* (pits = 20, scores = 34). Considering how the objective of this study was to provide a methodological approach for the study of tooth marks, the nature of the canid and second felid was not considered important, however will need to be rectified when a more complete study of FLK-West is performed.

Tooth mark samples were created in a controlled setting, while the bones contained flesh. For Hyena, Jaguar and Lion studies, these marks were obtained from the Cabárceno Park, Cantabria (Spain). Tooth pits for wolves were also obtained from Cabárceno, however tooth scores for wolves were sampled from a natural wolf site in Mount Campleo, near Sobrado Dos Montxes, Galicia. The natural wolf sample was produced by a range of between 10 and 15 wild individuals.

Tooth marks were only selected from long bone shafts (tibiae and radii). This can be explained for two main reasons; firstly diaphysis are denser than epiphyses, therefore teeth tend to penetrate less, and secondly, diaphysis present a higher survival rate, thus proving to be a more useful experimental

comparative framework for archaeological analogies. The use of tibiae and radii as anatomical elements for the experiments were conditioned by the workers of the natural parks who helped collect these samples.

The teeth marks used were selected according to their preservation and general conditions.

The origin and original description for each tooth mark sample can be found in:

- Gidna et al. (2013) for the case of lions.
- Domínguez-Rodrigo et al. (2015a) for the case of hyenas.
- Domínguez-Rodrigo et al. (2015a) for the case of jaguars.
- Yravedra et al. (2011, 2012, 2017c) for the case of wolves.

These samples have in turn been studied and published by multiple authors for similar purposes, including Aramendi et al. (2017a), Yravedra et al. (2017c) and most recently by Courtenay et al. (2019), included in Appendix 5, as a product of this Master's Thesis.

4.1.4. Percussion Marks

Percussion mark samples were taken from Yravedra et al. (2018). This sample consists of 39 percussion marks experimentally created using unmodified quartzite hammer stones. Only percussion marks not associated with microstriations were selected for this study, considering their higher equifinality when being compared with carnivore tooth marks. Quartzite was selected as the raw material considering its abundance in the archaeological record of the Olduvai Gorge. Nevertheless, previous experimentation has also proven how the raw material is not a conditioning factor in percussion mark morphology (Yravedra et al., 2018: S1).

4.2. Archaeological Sample: The Site of FLK-West

The FLK-West sample included a total of 8 bones that were selected for this study, presenting a total of 20 linear and 4 circular marks. The bones presented consist of 1 indeterminable bone, 1 humerus, 1 metatarsus, 1 tibia, 2 femora, 2 radii. The majority of these individuals are from Size 3 bovid individuals, minus the metatarsus which belongs to a *Sivatherium*. Femora contain the highest number of linear marks and also present the 4 circular marks included in this analysis. Tibiae on the other hand also present frequent linear BSMs followed by the metatarsus and finally the radius.

Previous taphonomic analysis of these remains have been able to highlight the important presence of fluvial rounding and abrasion of cortical surfaces (Yravedra et al., 2017a), while GMM studies of some of these cut mark cross sections have revealed quartzite to be the main raw material acting on these bones (*ibid*). Considering the depositional environment of FLK-West, these observations of fluvial abrasion are not surprising, however the effect this abrasion has had on the already present BSMs seems to have increased the level of equifinality present at this site (Courtenay, 2017).

Photos of the studied pieces can be found in Figures 4.5. and 4.6..

Figure 4.5. - Figure presenting the studied artefacts recovered from the archaeological site of FLK-West.

Figure 4.6. - Figure presenting the metatarsal of a Sivatherium recovered from the archaeological site of FLK-West.

4.3. Microscopy

Multiple different microscopic methods of processing taphonomic BSMs were used for this study. While most concerned the digital reconstruction of these marks for advanced statistical analysis, other microscopes and methods were also used for documentation. The objective of these approaches were to provide the highest level of documentation for the taphonomic register presented in the site of FLK-West, as well as the most efficient means of studying this. In total 3 different techniques were used for graphical documentation, the first being the Environmental Scanning Electron Microscope (ESEM), the second being the HIROX KH-8700 3D Digital Microscope (HIROX) and the last being standard photography. In addition to this, 3D reconstructions were produced using primarily the HIROX, to a lesser degree the DAVID Structured-Light SLS-2 Scanner and finally microphotogrammetry. Each of these techniques present their advantages and disadvantages, as will be explained throughout this Master's thesis.

4.3.1. Scanning Electron Microscopy

SEM is a non-destructive method that utilises electronic lenses and electron waves to provide high resolution images with a greater perception of texture than standard optic mircoscopes. The Environmental SEM (ESEM) used in this study is the FEI Quanta 600 Scanning Electron Microscope located in the *Servei de Recursos Ceintífics I Tècnics* of the *Universidad Rovira I Virgili*, Tarragona. The resolution of the microscope reaches up to 1.2 nm at high vacuum and 1.5 nm at low vacuum. This system utilises a voltage of 200 V to 30 kV. The magnification obtained using this equipment can reach up to 2,000,000x magnification, depending on the user interface and the LCD monitor used. This piece of equipment is additionally configured with a Noran energy X-Ray analyser that permits for chemical analysis of specific points as well as the mapping out of the chemical elements present.

The SEM is equipped with a motorized, tiltable and rotatable stage that can be used to adjust the position of the object under observation. The Focus range provided is from 2.5 – 99m while the FOV can reach up to approximately 17mm at the longest working distance.

4.3.2. 3D Digital Microscopy

3D Digital Microscopes use digital cameras, incorporated computer software and optical systems to produce the required image for observation and even 3D analysis. The 3D Digital Microscope using in this study was the HIROX KH-8700 3D Digital Microscope (Fig. 4.7.) located at the *Institut Catalá de Paleoecologia Humana I Evolució Social (IPHES)* of Tarragona. This microscope is equipped with a MXG-5000REZ triple objective revolving lens, with a magnification range from 35x to 5000x, with a field of view ranging from 8 mm to 0.06 mm at an operable distance of 3.5 mm – 10.0 mm (Table 4.1.). The microscope is accompanied by a high intensity LED light source that provides a temperature of 5700 k. This light source closely portrays daylight colour and produces the highest quality real-time images with no warm up time needed. The HIROX provides the possibility of alternating between coaxial and ring light.

Polarised filters can be added according to the analyst's needs.

The microscope is equipped with a built in compact CCD camera that projects images onto a high definition LCD 21.5" monitor. This system provides high intensity pixel reproduction displaying up to 16.77 million colours, utilising a contrast ratio of 1000:1 and brightness of 300 cd/m^2.

A combination of state of the art hardware and the Genex Engine graphics processor ensures that each image is displayed with maximum quality. The HIROX is also capable of producing 3D digital reconstructions using a combination of quick auto focus and depth synthesis functions. The HIROX also provides a tiling function that can be used to create mosaic reconstructions of complete surfaces in 3D.

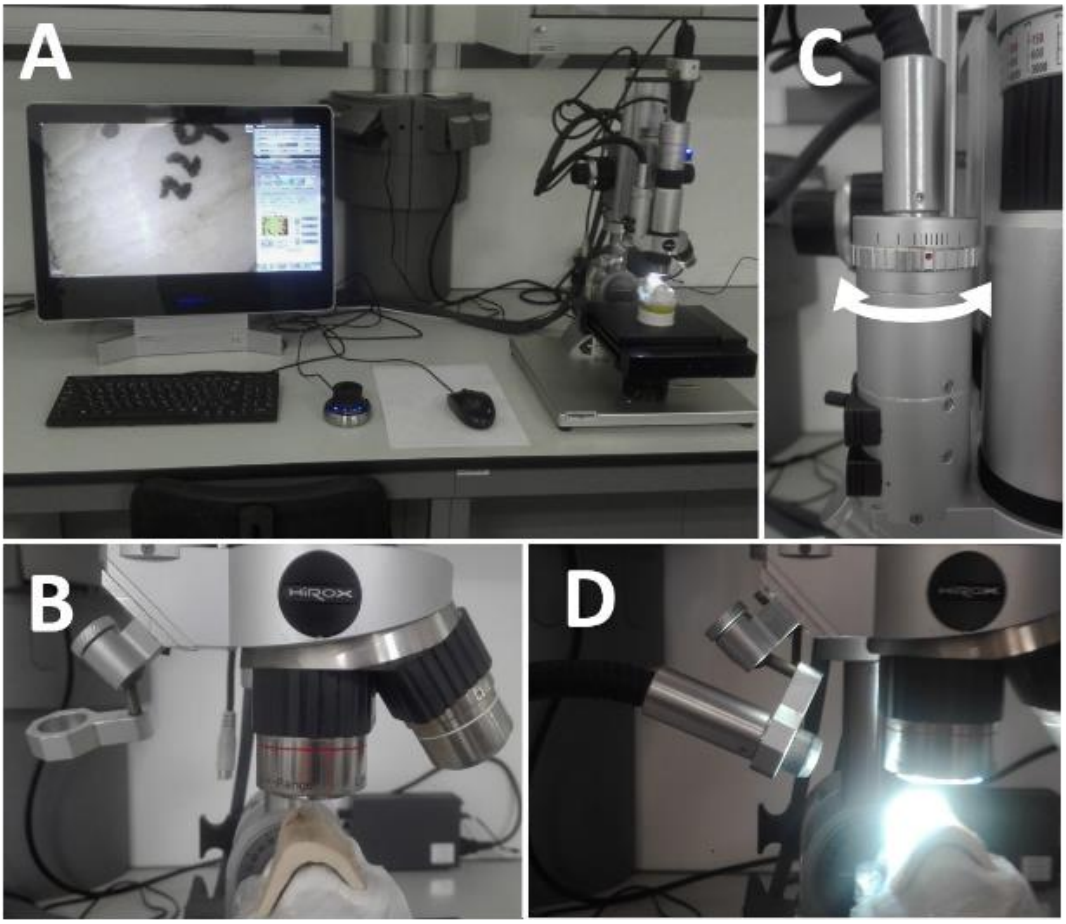


Figure 4.7. – (A) The HIROX KH-8700 3D Digital Microscope located at the IPHES lab. (B) The MXG-5000REZ Triple Objective Revolving Lens. (C) Lighting Position from above providing the option of combining ring (turning the wheel to the right) and coaxial (left) lighting conditions. (D) Lighting position from the side, using the movable adjustable light support.

Lens Magnification	Field of View	Depth of Field	Working Distance
35x - 250x	8.76 - 1.22 mm	0.72 - 0.72 mm	10 mm
140x -1000x	2.18 - 0.31 mm	0.09 - 0.007 mm	10 mm
700x - 5000x	0.44 - 0.06 mm	0.01 - 0.0007 mm	3.5 mm

Table 4.1. – Details regarding each of the lenses of the MXG-5000REZ Triple Objective Revolving Lens

4.3.3. Additional 3D Reconstructions

Two additional 3D reconstruction techniques were employed. The first consisted in microphotogrammetry while the second employed the use of the DAVID Structured-Light SLS-2 Scanner, located at the TIDOP research group in the University of Salamanca and obtained through the CAI Ciencias de la Tierra y Arqueometría department of the Complutense University of Madrid. These two techniques were carried out by a specialist (Miguel-Ángel Maté-González) considering how they were used on previously described experimental samples from other papers (Maté-González et al., 2017a; Aramendi et al., 2017a; Courtenay et al., 2017; Yravedra et al., 2017c, 2018).

Microphotogrammetry was performed using a CANON EOS 700D camera with a 60 mm macro lens. The protocol used was that specified by Maté-González et al. (2015). The camera was placed in automatic to simultaneously compute the interior and exterior parameters for the camera and the lens. 9-10 photos were taken for each mark. The number of photos varying depending on the geometry of the bone and shape of the mark (Aramendi et al., 2017a, Yravedra et al., 2017c). Once the photographs had been taken, they were processed in the open source photogrammetric reconstruction software GRAPHOS (integrated PHOtogrammetric Suite) (González-Aguilera et al., 2016).

Digital reconstructions via the DAVID SLS-2 consist in the use of a DAVID USB CMOS Monochrome camera, an ACER K132 projector and a calibration marker board. This piece of equipment generates a point density of up to 1.2 million while creating high resolution 3D images in less than 1 min.

The first phase of the scanning process consists in calibration of the equipment, performed by placing the camera alongside the projector, both facing towards the calibration marker board at an angle between 15º and 25º. The projector is then used to project a standardised image over the marker board. In this case a 15 mm scaled calibration pattern was used so that the DAVID software could adjust the camera's exposure accordingly to ensure optimal focus of all instruments. During this process of calibration, neither the calibration board nor the camera/projector can be moved, ensuring they are fixed and stable throughout the entire process.

The second phase of the scanning process consists in replacing the calibration board with the bone to be scanned. The scan is then executed, producing a high resolution 3D image that can be exported and processed in a number of different computer softwares.

Previous efforts into understanding the methodological limitations and advantages of these techniques have already been able to conclude that no statistical differences are presented when comparing microphotogrammetric models with scanned models (Maté-González et al., 2017c).

One other surface scanner was considered for this study, being that of the Breukman Smartscan located in the IPHES laboratory, Tarragona. The resolution this scanner presents, however, is dependent on the lenses used for capturing 3D data, and the lenses currently available in the IPHES lab are orientated more towards the digitalisation of large objects for archaeological and paleontological studies, such as those concerning entire skull morphologies or lithic implements, for example. The consequent resolution for the processing of microscopic taphonomic data, therefore, did not match the requirements of this study (Fig. 4.8.).

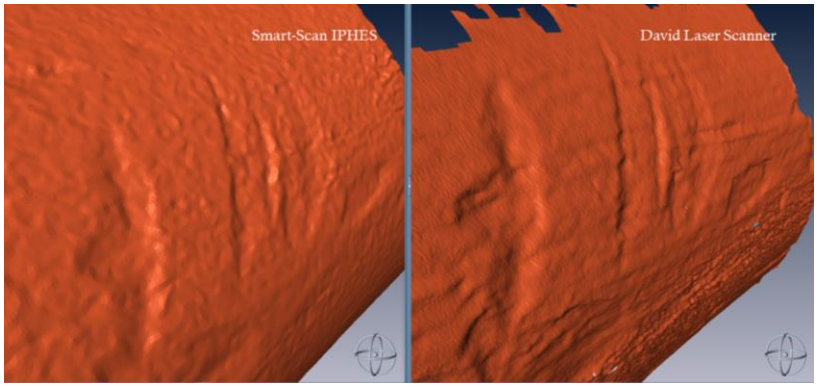


Figure 4.8. – Comparison of the resolution produced by the Smart-Scan (left) and the DAVID SLS-2 (right) for processing taphonomic traces.

4.3.4. Processing 2D and 3D Reconstructions

The processing of 3D images were performed in multiple computer software applications. Firstly, different marks processed using microphotogrammetry and structured light laser scanning were imported into the Global Mapper v.15 software (Blue Marble Geographics, 2019). Here marks could be treated, extracting cross sections at each incisions' midpoint (ensuring that the section lies in between 30 and 70% of the marks' total length) (Maté-González et al. 2015). Cross sections were then imported into tpsDig2 v.2.29 (Rohlf, 2017), where 2D landmarks could be positioned and exported as a .tps file. For 3D GMM models, however, the Amira 5.0 software (Visualization Sciences Group, USA) was used.

In the case of the HIROX, cut mark profiles were extracted as normal, and then imported into tpsDig2 v.2.29 for processing. 3D landmarks, however, were not as simple. In order to collect this landmark data, the position of each point was recorded through a series of measurements. These measurements were provided by the HIROX's own internal software, thus reducing accurate results with little to no human error. The HIROX's own self-calibration select sensor automatically configures and applies the appropriate lens settings according to the lens and magnification used for measurements of up to 1µm accuracy. Through this, the 'XY-Width' function was used to plot the position of the landmark across a 2D graph for each point, followed by the 'point height' function to calculate the position of this landmark on the z-axis along a 3D plot. Landmarks were recorded into a database that could be easily formatted for further statistical analysis.

An example of this process to process cut marks is described in Fig. 4.9., while tooth pits can be seen in Fig. 4.10..

Of the 4 different experimental samples, the HIROX was used for the majority of digital reconstructions. This includes both trampling and cut mark samples, minus the case of cut marks produced by basalt and different tool types from Courtenay et al. (2017). The HIROX was also used to process all of the taphonomic marks presented in the FLK-West material. The DAVID SLS-2, on the other hand, was used primarily to process tooth pit and percussion mark samples. Microphotogrammetry was finally used to process tooth scores. The use of different techniques for different samples was due to the availability of the different pieces of equipment when attaining and processing each sample. As shown by Maté-González et al. (2017a, c), however, all pieces of equipment have not been found to present significant statistical differences between samples.

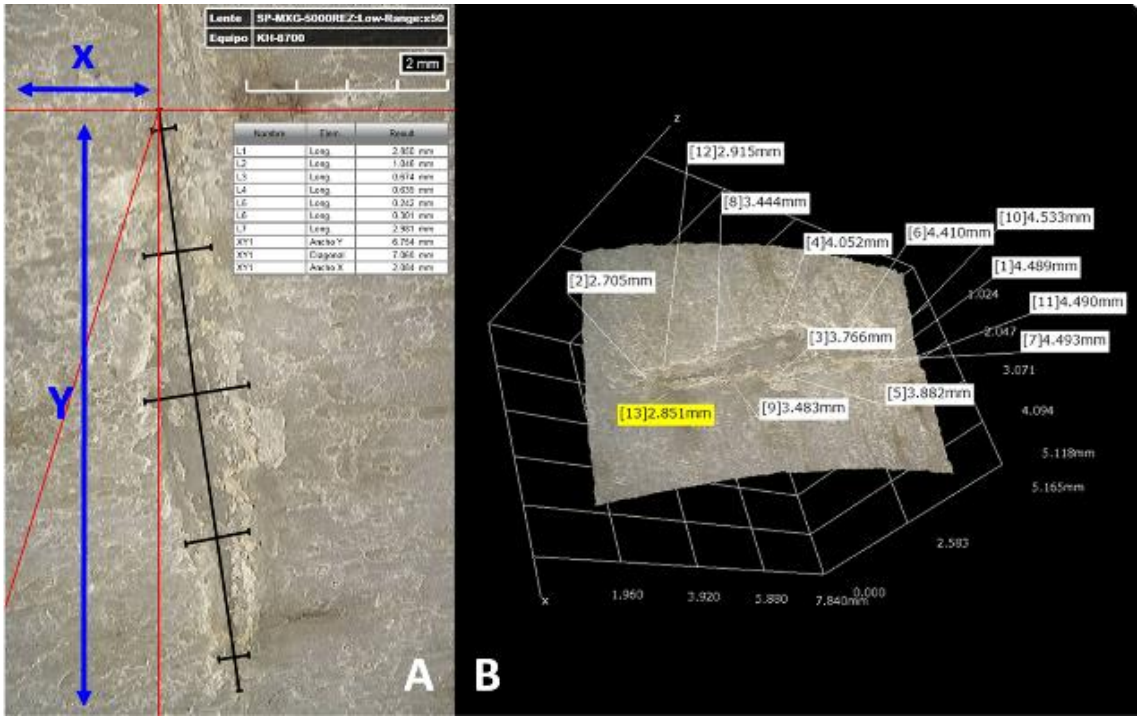


Figure 4.9. – Image to describe how measurements are taken to record landmark data for linear incisions. (A) x and y axes; (B) z axis.

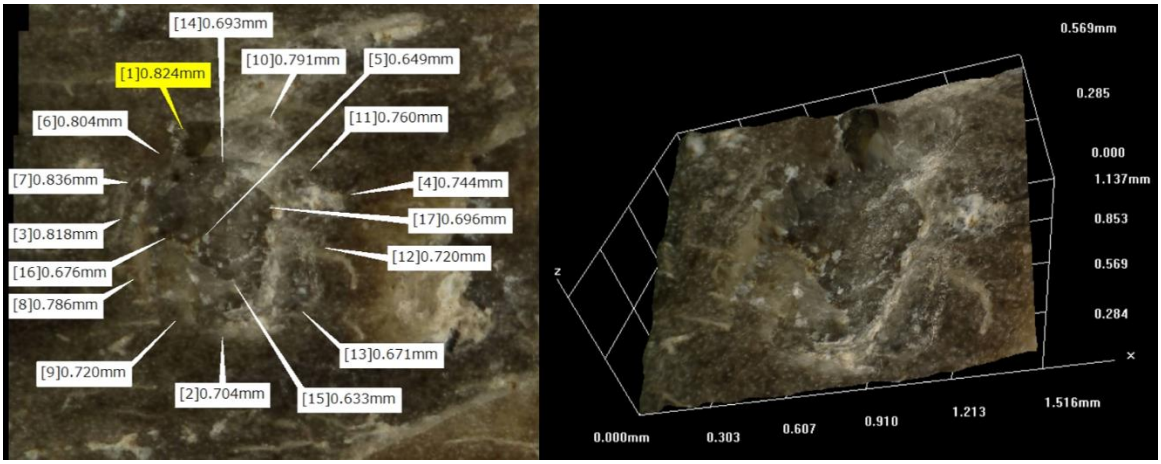


Figure 4.10. – Image presenting reconstruction of circular depressions and measurements of depth to record z axis coordinates

4.3.4. Comparing Methods

Before taphonomic analysis could be performed, statistical comparisons of the HIROX KH-8700 3D Digital Microscope and the DAVID Structured-Light SLS-2 Scanner had to be performed to test that marks studied with these techniques are comparable. This was also considered a necessity in order to understand the variation that the HIROX microscope may produce under the different lighting conditions.

For this part of the study, a total of 5 different lighting conditions were initially considered for the HIROX; from the left of the mark, from the right of the mark, a fixed position above the mark using a ring light, a fixed position above the mark using a combination of ring and coaxial light and finally the use of two light sources from either side of the mark. The original intention of using two light sources aimed at reducing the amount of shadow cast over the incision and thus provide sufficient illumination for the entire mark. The strength of the HIROX's own light source, however, presented multiple problems. Different types of secondary light sources were trailed and tested, however, none of these lights were able to meet the HIROX's intensity. Because of this, the 5th position was discarded.

To ensure that no other variables were producing statistical noise in our analysis, neither the bone nor the platform could be moved of altered during the digital reconstruction process. This ensures that the only changing variable in this analysis was the positioning of the light.

Cut mark sections were processed using the mid-range lens at 600x magnification, while 3D models were constructed for the entire mark at approximately 50 - 100x magnification.

A random selection of 5 cut marks were reconstructed using all 4 lighting positions. These cut marks came from previously studied material by Courtenay et al. (2017), all produced using quartzite simple flakes on a *suid* femur diaphysis. These were then processed using both Courtenay et al. (2017) and Maté-González et al. (2015)'s GMM methodological approaches that are explained on detail in Fig. 2.22 on Page 42. This study enabled the evaluation of morphological differences that could be produced by the 3D reconstruction method, thus testing the reliability of the HIROX for 3D GMMs.

Additional notes on the amount of photos taken per tile, the number of tiles used to reconstruct the full incision, the time taken to reconstruct each mark and the time taken for the analyst to process the mark were taken. Alongside this, the number of pixels, distance between the lowermost and uppermost point of the lens were also noted, all used later for statistical analyses.

Once the optimal lighting position had been statistically established, this was then used for all applications of this technique to 3D reconstructions.

4.4. Advanced Statistics

This study employed the use of six main computer software for data science applications (Fig. 4.11.).

Most statistical analysis were carried out in the free software R, x64 v.3.5.1. (www.rproject.org, Core-Team, 2018). R is an open access statistical computing software that employs its own programming language and syntax for efficient computation and data mining. R implements a wide number of different techniques to carry out each study, and uses multiple libraries and packages that can be downloaded easily and for free with just one simple line of code. While R uses its own syntax, users of programmes such as C++, Java and Python can still familiarise themselves quickly with the coding language employed. In some features, such as in DL, different software can be linked to provide an efficient data analysis experience. Finally, R offers multiple different platforms that can be used to share research and provide support for users, including an open access and refereed indexed statistical journal known as *The R Journal* (https://journal.r-project.org).

Additionally, RStudio, x64 v.1.1.463 (www.rstudio.com, RStudio Team, 2015) was used to prepare all the R code employed in this study. RStudio is another free and open source software that provides an extension to the R environment and workspace. In RStudio, code can be prepared, trialled, debugged and

executed for a more efficient programming experience using R syntax. Once the code is ready and finished, it can be saved externally and executed by any computer with R installed. An important feature of RStudio is its' easily readable and modifiable interface, with colour-coded syntax that makes writing and refining code much easier.

R and RStudio are exceptionally powerful computational systems for advanced statistics, requiring little processing power to efficiently provide complex mathematical computations. Most calculations can last a matter of seconds and can be carried out using short lines of coding.

For DL, the free statistical software Python was used, x64 v.3.7. (http://www.python.org, Python Software Foundation, 2018). Python is the most popular application for programming of DL systems. Python employs the use of syntax similar to C++ and java applications and thus proving to be readily usable by most users. Python, much like R, employs the use of multiple packages that are easily manageable when performing different tasks or executing long blocks of coding. Python additionally offers multiple platforms for support, using community-based development programmes to ensure that analysts of any level are able to use this software.

Package management, environment handling and preparation for AI practices employed the Anaconda, x64 v.1.9.2. (Anaconda Software Distribution, 2016), data science platform. This is a highly efficient open source platform that helps with the distribution of Python and R programming languages in the data science community. Anaconda is used to prepare the computers virtual environment for analytical activities such as ML. This platform is also highly useful for managing Python in DL. While it is not a necessity nor is it essential for ML, Anaconda provides the analyst with all the material necessary to prepare the code that will be used for ML or DL. In this study, Anaconda and the Conda cross-platform were used to manage packages, libraries and working environments for all AI applications.

Debugging and initial development of DL python code was carried out using Jupyter Notebook, x64 v. 5.7.8 (Kluyver et al., 2016). Jupyter is an open-access browser based software that provides computing for multiple platforms such as Python. This web application provides a free service with a highly interactive interface that allows the user to execute large chunks or code bit by bit. The Jupyter interface also has a strong community based approach to sharing code and helping analysts prepare as well as improve their applications. For this study, Jupyter's debugging options were used making the most of the ability to debug chunks of code at a time, thus identifying possible programming errors early on before the code could be finalised.

Finally, Visual Studio Code (VS Code) software, x64 v. 1.32.3. (code.visualstudio.com, Microsoft, 2015), was used for final tweaking of the design and last minute preparations of DL code for Python. VS Code is a user friendly interface and editor for the preparation of codes for any programming language. VS Code allows users of Python, R, C++, Java and many others to prepare their script prior to execution. The use of VS Code was preferred for this study considering the colour coded syntax, highlighting and debugging options that are highly useful for saving time and reducing errors when trying to execute new codes directly in the Python shell.

Figure 4.11. - Software and Programmes used for AI applications and statistics. From left to right: (Top) R and Python, (Bottom) Anaconda, RStudio, Visual Studio Code and Jupyter Notebook.

Data science and statistical applications were carried out on an ASUS X550VX (R510VX-DM607) personal laptop, running in Windows 10 and with an Intel® Core™ i5-6300HQ processor. The system runs with a 64-bit operating system, has a Central Processing Unit (CPU) of 2.30 GHz and an available Random Access Memory (RAM) of 8 GB. The hardware additionally presents an Intel® HD Graphics 530 card, proving highly efficient for working with and visualising 3D models. Additionally, the laptop has Java x64 v.1.8.0_151 installed in order to run R and other software.

Data visualisation was either done using internal packages for each software, such as ggplot2 for R (Wickham, 2016) and Matplotlib for Python (Hunter et al., 2019), with any additional editing carried out in the ClipStudio Paint software, x32 v. 1.8.4 (CELSYS inc., 2014) using a WACOM graphics tablet.

4.4.1. Taphonomic Analyses

The taphonomic part of this study included a number of classical taphonomic analyses. These were carried out to compare the new methodological approaches with older ones when applied to the archaeological register. Additionally, this approach intended to provide a further description of the trampling marks and cut marks studied.

For cut and trampling marks, a database was created in Microsoft Access recording each variable described by Domínguez-Rodrigo et al. (2009a). These include a total of 16 variables including:

1. Groove Trajectory (Straight, Curvy, Sinuous)
2. Barb (Present, Absent)
3. Groove Shape (\/, _/)
4. Groove Symmetry (Symmetrical, Asymmetrical)
5. Shoulder Effect (Present, Absent)
6. Flaking on Shoulder (Present, Absent)
7. Extent of Flaking (Less or more than 30% of the mark)
8. Internal Microstriations (Present, Absent)

9. Trajectory of Microstriations (Continuous, Discontinuous)
10. Shape of Microstriations (Straight, Irregular)
11. Location of Microstriations (Walls, Bottom, Both)
12. Mark Orientation (Oblique, Parallel, Perpendicular)
13. Associated Microabrasions (Present, Absent)
14. Overlapping Striae (Present, Absent)
15. Length of Main Groove (mm)
16. Number of Conspicuous Grooves per Specimen (Int.)

In order to consider purely qualitative variables, and thus compare methods, the "length of the main groove" and "grooves per specimen" were excluded from this analysis. Additionally, in some cases, the variable "Mark Orientation" could not be recorded or compared with other samples, therefore to ensure replicability this variable was excluded.

Once frequency tables of each qualitative feature have been recorded for trampling marks as well as cut marks, a Multiple Correspondence Analysis (MCA) of these frequencies can be performed. MCAs have proven effective when comparing multiple tables, as demonstrated by Moclan et al. (2018) and Pineda et al. (2014). MCA was performed in the 'ca' R package (Nenadic and Greenacre, 2007).

To try and remain objective, features such as the shape of the groove were confirmed withdrawing 2D cross-sections of the grooves using the HIROX. Variables such as the extent of flaking were measured and the use of high magnifications aided the rigorous examination of each mark to confirm the presence and nature of microstriations, for example.

Frequency tables produced were additionally analysed using two-sample T-tests and Chi-Squared tests to calculate the significance of observed frequencies. These are carried out in the package 'MASS' package (Venables and Ripley, 2002). This method was equally used to calculate the significance of trampling mark frequencies under the different sedimentological and taphonomic conditions.

To support this study, additional metric analyses of the cut and trampling marks were performed to truly understand the nature of these different linear traces.

For this part of the study measurements were taken including the length, width and depth of each linear mark. Measurements for width and length were calculated using landmark coordinate data derived from the 3D models. This is done taking the landmarks that mark the beginning, middle and end of each incision, and then using the formula:

$$d = \sqrt{x^2 + y^2 + z^2}$$

where *x*, *y* and *z* are the coordinates in question. The inclusion of the central landmark was used to factor in the curvature of the bone in this calculation. This was then carried out to calculate depth and width of each mark and store this data in a separate database.

To ensure that other analysts are able to replicate the bidimensional statistical analyses performed here, each was additionally retaken using a digital calliper. Degrees of differences between the samples were calculated using Analysis of Variance (ANOVA). Once these bidimensional variables were obtained, the samples were tested using the Spearman Coefficient for correlations. This was followed by the performance of multiple regressions to see how the variables could be used to distinguish between the observed groups.

Regressions are a means of calculating the relationships between variables. Standard regression techniques include the use of one or more variables to see how they affect other variables in the data set. The mathematical means of describing this relies heavily on the "line of best fit", which considers a simple equation described through:

$$y = mx + b$$

where *y* is the dependent variable, *m* is the point where the regression line intercepts the y axis, *b* is the slope coefficient and *x* is the independent variable. The line of best fit (conditioned by values *m* and *b*) is plotted where all the total prediction errors are as small as possible.

Many types of regressions exist depending on the data being used, however, the best means of finding the best "fit" is through trial and error, comparing each regression's performance and representation of the data. This evaluation essentially quantifies the strength of our correlation, and can even be employed when classifying or predicting data based on these variables.

Regressions are very easy to perform and even simple software such as Microsoft excel can be used to carry one out. Nevertheless, the power of R's internal algorithms for calculating and comparing formulas is preferred for this particular case.

The first type of regression performed was a linear regression. This standard calculation observes the difference between two variables, one independent and one dependent variable. The calculation assesses how the dependent variable is effected by change in the independent variable. The concept of linearity in this equation assumes that either all the values are constant, or that the independent variable is being directly affected by a hidden parameter.

The second type of regression used considers multiple independent variables, known as a multiple linear regression, thus being able to plot up to 3 dimensions or over. In many cases, these type of calculations are impossible to visualise and handle. To overcome this step-regression models are used to calculate the significance of each variable and adjust accordingly.

In the case where values are not constant, and therefore non-linear, linear models are likely to produce a much larger margin of error when calculating the line of best fit. These variations can be explained by the nature of *b* in our original calculation. Logistic models, for example, are a type of binomial regression that employ the logarithm function in order to correct the line of best fit to the variables presented. In many cases, this model can use sigmoid functions rather than logarithmic functions to provide the lowest predicted error. The nature of the transformation used is dependent on the analyst and should be explored when testing different types of data.

Other types of non-linear regressions include binomial statistics, which fall closer to multiple linear regressions in nature, considering the wide range of variables that can essentially affect their distribution. This calculation is best represented through the adaptation of the previous calculation to include the value *c* which represents an unobserved variable yet to be defined and discovered:

$$y = mx + b + c$$

If we are to plug this function into a multiple variable environment, the calculation becomes:

$$y = b + m_n x + c$$

This can reach great complexities depending on the *n*th number of variables involved.

The final model employed was the *robust* regression model. These use a different optimisation metrics to find the line of best fit, avoiding the use of underlying mathematical assumptions. Such assumptions are highly evident in other regression models, where the nature of the data's distribution regarding homogeneity and linearity could be a conditioning factor in the test's final results. Through this method, the calculation ensures that the error values predicted are carefully selected so as not to effect the entire distribution, considering and consequently excluding the presence of obvious outliers and anomalies.

Evaluation of regression models are typically performed using two main parameters; the predicted error values,

otherwise known as residuals and the Akaike Information Criterion (AIC). Residual values can be described as the distance between each point and the line of best fit. These values can be plotted across multiple graphs and assessed in reference to the regression line as well as the overall variability presented by the sample. AIC values are simple to interpret, whereby the closest value to 0 (whether this be positive or negative) presents the best fit to the overall variation presented by the sample. AIC is a function provided by the stats package (www.rproject.org, Core-Team, 2018), while residuals can be assessed in any of R's packages that include regression analysis.

The final statistical test employed was the "two one-sided" equivalence test (TOST) from the TOSTER library (Cohen, 1988; Lakens, 2017). This test is used instead of Multivariate Analysis (MANOVA) tests depending on the null hypothesis being tested. TOST provides an assessment of the magnitude of differences through calculating the significance of overlapping in samples, according to Cohen's d statistic. This procedure originates from pharmacological studies, and is used mostly as a means of assessing similarities between pharmaceutical products (Hauk and Anderson, 1984). The TOST test produces graphs and values that evaluate the equivalence bounds of a sample based on a specified effect size (Schurimann, 1987). The *p* value produce assesses the degree of similarities, thus identifying $p > 0.05$ to mean significant similarities between groups, as opposed to MANOVA where $p > 0.05$ indicates significant differences. The different formulae and calculations that can be used in TOST are explained in detail in Lakens (2017).

4.4.2. Geometric Morphometrics

For this study, 4 different GMM landmark configurations were used:

1. Courtenay et al. (2017)'s 3D 13 landmark model was the main model used in this study (Fig. 2.22, Page 42). This consists of a mixture of Type II and Type III landmarks that are mapped out across the inside and the outside of linear taphonomic marks, namely cut marks.
2. Maté-González et al. (2015)'s 2D 7 landmark model for the study of cut mark cross-sections was also used (Fig. 2.22, Page 42). This consists of a mixture of Type II and Type III landmarks that are mapped out along the cross section, adapting Bello and Soligo (2008)'s method to quantifying cut mark cross-section dimensions.
3. Yravedra et al. (2017c)'s 2D 7 landmark model for the study of tooth score cross-sections (Fig. 2.22, Page 42). Similar to Maté-González et al. (2015)'s model, this consists in a mixture of Type II and III landmarks.
4. Aramendi et al. (2017a)'s 3D 17 landmark model for the study of tooth pits, adapted by Yravedra et al. (2018) to include all circular depressions such as percussion pits (Fig. 2.23., Page 42). Again, this configuration consists in Type II and Type III landmarks mapped out across the inside and the outside of the depression.

Landmark coordinates obtained via the 3D reconstructions and 3D derived 2D cross-sections were edited and imported into R. The editing process transforms each .txt and .tps file containing landmark information into a single

.txt file edited into the Morfologika (*.txt) format. Morfologika files are a standardised format used specifically for GMMs, following the guided instructions of O'Higgins and Jonhnson (1988) and O'Higgins and Jones (2006). These files contain all the landmark coordinate information for each individual associated with their respected labels and label-values. These files can be read into R using the `read.morfologika()` function of the 'geomorph' library (Adams et al., 2017).

Once imported into R, an orthogonal tangent projection was used to normalise the data through GPA for further multivariate statistical analyses (Dryden and Mardia, 1998). A Principal Components Analysis (PCA) was then derived from this data in both form and shape space in order to reduce the large sets of variables into fewer dimensions with a greater representation of overall variance. These scores (PC scores) produced through this analysis can then be extracted for further statistical analaysis. Changes across each score were additionally visualised using transformation grids (Bookstein, 1989).

Once the PC scores for each sample were extracted, statistical significance and degrees of variance were tested using multiple different techniques.

First the homogeneity of each sample was evaluated using the Shapiro test. Depending on the normality of the represented sample, one of two different Multivariante Statistical Analyses (MANOVA) was performed to test for similarities and differences between groups. In the case of a normal distribution, a MANOVA test was performed using the 'stats' package (www.rproject.org, Core-Team, 2018). If the Shapiro results tested abnormally, then a pairwise permutation MANOVA test was performed in the 'RVAideMemoire' package (Hervé, 2018).

The use of different formulae to calculate statistical degrees of variation allows for the analyst to adapt according to the nature of the sample being tested. Many statistical tests contain mathematical assumptions that need to be addressed when choosing the right test. For this, two different MANOVA tests are used depending on the normality of the sample. The Hotelling-Lawley calculation is a powerful calculation that can be used to determine significance levels (Lawley, 1938; Hotelling, 1951; Coombs, 1996), the trace statistic (T_g^2) employed can be defined as follows:

$$T_g^2 = e \sum_{j=1}^{s} \emptyset_j$$

where e is the degree of freedom for the error matrix and s is min(p,h), p being the probability and h being the hypothesis matrix (Hintze, 2007:415-2). The Wilks Lambda ($\Lambda_{p,h,e}$) on the other hand is described as:

$$\Lambda_{p,h,e} = \prod_{j=1}^{p} (1 - \theta_j)$$

where p, h and e are the same as the case described before (Rao, 1951; Hintze, 2007). Each presents its particularities that confront the sample under study considering the sample size and within-group homogeneity to reduce statistical errors (Coombs, 1996). For this study, the Hotelling-Lawley trace calculation was used for normal distributions, while the Wilks Lambda calculation was preferred for non-normal distributions.

Following the MANOVA tests, a Canonical Variance Analysis (CVA) was performed, following standard GMM protocol in typical studies of this nature. CVAs are common practice in comparative studies of samples, a useful procedure to observe the nature of inter-group variance. This procedure consists in a transformation of the raw data, generally PCA, whereby the pooled within-group dispersion are manipulated in a scaling process, thus standardising within-group variance, and finally rotating the axes to provide a redrawn CVA graph (Albrecht, 1992; Klingenberg and Monteiro, 2005). From these standardised values, distances can be calculated

between the groups with permutated p values from the pooled within-group covariance matrices, calculating a degree of separation between samples and the inter-group variances present.

The two main distances calculated through this approach are the Procrustes and Mahalanobis distances (Klingenberg and Monteiro, 2005). Some specialists have used this approach as a means of computing classification probabilities (Ross et al., 2010), however, great care should be taken with this approach concerning sample size especially. Because CVA works with in-group variation before the standardisation process is performed, all the samples of a single label will be grouped together (Albrecht, 1992). Therefore, if we were to use this technique to classify an unknown sample of mixed archaeological trampling and cut marks, the test directly assumes that all the individuals within this group are the same. Through this, the entire group will be classified as one or the other, without possibilities to separate the unknown samples. Likewise, marks that neither belong to one group nor the other, will still be misclassified according to their association with the rest of the sample. Because of this, CVA will *not* be used in this case as a classification method, yet will be used solely to evaluate the degree of inter-group variations in our experimental samples.

When included, CVA was performed in the 'shapes' package (Dryden, 2018), while Mahalanobis and Procrustes distances were calculated with the combined usage of the 'stats' package (www.rproject.org, Core-Team, 2018).

The final classical GMM approach consists in defining a classification model, most typically a Linear Discriminant Analysis (LDA). LDA assumes linearity among groups to find the best division between each sample. These classifications are formed generally using pairs of groups, and can then be summarised in the final model. The largest problem with LDA are the assumptions the test makes when analysing data, coming to the same problem as MANOVA in which multivariate normality, homogeneity, linearity and independence of the variables affects its performance. Nevertheless, despite these assumptions, LDA is assumed to be a very robust classification method in most cases. Because of these assumptions, the calculation used can vary greatly.

The most common LDA calculation used in GMM is that of Lachenbruch (1967) updated by Dunne and Stone (1993). This function, for example, is that which is included in Klingenberg (2011)'s MorphoJ software. This calculation follows the leave-one-out cross-validation discrimination function, whereby each value in turn is excluded, and then used to reclassify. The degree of errors is calculated when the omitted observation is misclassified. This particular calculation is represented by multiple mathematical formulae that can be referred to in Lachenbruch (1967).

For GMM LDA, the 'MASS' package (Venables and Ripley, 2002) was used.

4.4.3. Machine Learning

Two main types of ML algorithms were used. The algorithm used was dependent on the data being mined and the objective at hand.

For in group variation and pattern recognition problems, unsupervised algorithms were used. These algorithms can vary in complexity, ranging from standard clustering models (Kuhn and Johnson, 2013; Lantz, 2013), to DL autoencoders, that are based off of Neural Network architectures (Chollet, 2017; Géron, 2017; Paterson and Gibson, 2017; Patel, 2019). In this study, the K-means supervised model was used in the 'stats' R package (www.rproject.org, Core-Team, 2018). This clustering model works by taking an unlabelled

data set and creating new values from the given data, solving via various pattern recognition tests the optimum grouping of samples. This technique eventually creates a new variable which is considered the clustered label. This model is easy to train and is highly flexible to the data provided (Lantz, 2013), yet it is not as sophisticated as a DL model. The other disadvantage of this approach lies in how the algorithm requires an original input generally estimating the number of clusters to be expected. Nevertheless, cluster evaluation and optimisation metrics are available to test for the best fit on the data, reducing the subjectivity observed.

The most common K-means clustering technique employs the Nearest Neighbour algorithm that essentially calculates the cluster groups according to distance. The algorithm begins by sorting through the provided feature space, generating *k* points that are considered cluster centres. In most cases the Euclidean distance is employed to calculate this feature (Lantz, 2013). Euclidean distance is defined as:

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_z)^2}$$

where *x* and *y* are the coordinates of each point. Once the distances have been established, the model sorts through the possible clustering samples to finally provide a final output of group labels. The evaluation of the final cluster groups can be evaluated performing an elbow method test. This technique employs a looped function whereby the test is repeated multiple times using different *k* values, the algorithm then assesses the within group heterogeneity to identify the optimum clustering of samples. The concept of the *elbow joint* describes the point where no additional k-groups are able to produce significant changes in representation of within-group homogeneity or heterogeneity beyond a certain point. This concept is adapted from linear algebra and calculus, essentially defining the optimum grouping for the unsupervised method provided (Lantz, 2013).

The R code designed to carry out this function and produce a loop that calculates the optimum grouping is shown in Figure 4.12.. This code is essentially an adaptation of the mathematical formula described by Thorndike (1953). For the presented code to work, the only libraries necessary are the 'stats' package in R (www.rproject.org, Core-Team, 2018).

The elbow method, however, can be arguably subjective at times. In order to overcome this, the *Kneedle* algorithm in the 'Kneed' Python library (Satopa et al., 2017) was used to calculate and detect the point of maximum curvature, thus objectively establishing the elbow point mathematically. In this study, the additional use of a Nearest Neighbour (described in detail on Page 93) calculation was used to see if the two groups would eventually be reclassified as each other, and thus detect the degree of homogeneity through confusion matrices.

```
v<-apply(data,k,var) # where k is the k val
cs<-(nrow(data)-1)*sum(v)
for (i in 2:15) {
  cluster_x = kmeans(data, centers = i)
  cs[i] = sum(cluster_x$withinss)
} ; plot(1:15, cs, type = "b")
```

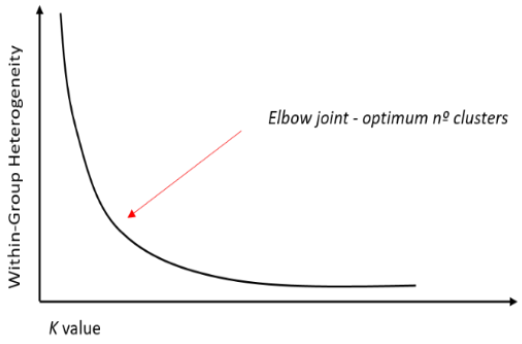


Figure 4.12. - R Code designed and used to create an Elbow-Joint graph where the optimal number of cluster groupings can be depicted graphically bellow

For classification problems, supervised learning algorithms were used. Mathematically, this method consists in providing an input (X) and the corresponding label set (Y) to find a means of reaching an output. The labelled set provides a series of examples that correspond to particular features in X. This data set can be denoted as follows:

$$S = \{(x_1, y_1), \dots, (x_n, y_n)\} \subseteq X \times Y$$

The output we intend to obtain with supervised techniques is described as a prediction rule or hypothesis ($h : X \rightarrow Y$), therefore, given a new X value, we intend to find the associated Y value. The supervised learning technique utilises performance metrics to evaluate the errors produced and optimise itself. The performance metric can vary in complexity and can also be conditioned by the type of algorithm being used.

From the different metrics evaluated, the analyst as well as the algorithm itself is able to update the parameters to enhance the model's performance, thus optimising the model we use for classification. Nevertheless, there are multiple different error functions and different possible algorithms that can be applied to achieve our classification goals and adjust each function. When using ML as a classifier, we need to consider the different algorithms available, test their performance on our particular data set, and finally use the best performing model to classify the unknown samples (Kuhn and Johnson, 2013).

The different Machine Learning algorithms employed in this study are as follows.

Neural Networks (NNET). These are inspired by human brain activity, using computed calculations to mimic brain activity patterns (Chollet and Allaire, 2017). Basic ML NNETs calculate relationships between the input and the output signal via a series of interconnected *neurons* or *nodes* to calculate the most efficient means of solving a mathematical problem. Just like in biological neural networks, an NNET established links between each neuron with weighted dendrites, a discriminant function that identifies the importance of variables and is used to help classify the inputted value (Lantz, 2013). This function can be described mathematically as follows:

$$y(x) = f\left(\sum_{i=1}^{n} w_i x_i\right)$$

where w are the weights of each n input (x). The final $f(x)$ activation function defines the resulting signal $y(x)$, which is passed on to the next neuron (Gurney, 1997; Heaton, 2012). The internal architecture of the NNET can vary in complexity, from simple ML NNETs to highly complex DL NNETs that take high levels of expertise to generate (Chollet and Allaire, 2017), as seen in the Deep Learning part of this study.

The performance of the NNET can be optimised by modifying internal parameters of the algorithm's structure, including activation thresholds, the number of layers involved, the number of nodes in each layer, and the final training of the network by tuning the weighted parameters via the *loss* function (Kuhn and Johnson, 2013; Chollet and Allaire, 2017).

The *loss* function and weight parameters are adjusted by the algorithm itself, whereby the degree of error produced during the training process is used to calculate and adjust the weights through the successive layers of nodes (Günther and Fritsch, 2010).

For this study, an ML NNET was constructed using the 'nnet' package (Ripley and Venables, 2016). The tuning of the model was performed in the 'e1071' R Package (Meyer et al., 2018).

Support Vector Machines (SVM). SVMs map out the input vectors into a non-linear high dimensional feature space, using hyperplanes to calculate the degree of separation between samples (Cortes and Vapnik, 1995). The feature space used in SVMs

is defined by a *kernel* function (Lantz, 2013). This particular approach is highly useful for instances where the data presented is non-linear. In most linear tests such as LDA, the linearity of a model can greatly affect the parametric nature of calculations, causing higher percentages of misclassification with regards to the line of best fit.

Kernel functions employ multiple calculations that essentially map out the data into a new space "domain", where linearity can be induced on these non-linear samples. In this study the kernel employed is the radial kernel, otherwise known as the Radial Basis Function (RBF), a kernel that performs well on most types of data and reaches the complexity of NNETs in performance (Müller, 2012; Lantz, 2013). This particular algorithm employs a series of complex mathematical formulae which are efficient for most training sets, regardless of the degree of overlapping in many cases (Bishop, 2006), reaching high accuracies independent of statistical noise in the sample.

The consequent hyperplane that SVM maps out to divide samples can be generally summarised to have *n*-1 dimensions, where *n* is the number of variables, therefore if 10 variables are present, the desired feature plane will have 9 dimensions to separate these samples.

SVMs are tuned similarly to NNETs using a *loss* function containing the residuals which tries to reduce overfitting. SVM was carried out using the "e1071" package for both training and tuning of the model (Meyer et al., 2018).

ML Decision Trees. Decision trees are similar to standard decision trees, with the additional complexity of a ML based approach to training the predictive model. Decision trees build models through recursive partitioning of the data set that acts much like a flowchart, where each branch of the tree is conditioned by a weighted value or variable that directs the input signal on to the next branch. This eventually reaches a final terminal or leaf node that provides the predicted output (Breiman et al., 1984; Quinlan, 1992). The main difference between ML trained models and standard decision trees lies in the adaptability to ML techniques of selecting and adjusting the internal architecture to select the variables that maximise performance accuracy (Wei and Chiu, 2015).

ML trees preform very highly on qualitative variables, however they are still adaptable to numeric data considering multiple regressions performed in their creation. The algorithm progressively partitions data according to the variables that will provide the greatest outcome. Common techniques in this process are the Standard Deviation Reduction process (SDR) which uses the following formula:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \cdot sd(T_i)$$

sd(T) refers to the standard deviation values of the set, while the formula in general measures the reduction in standard deviation from the original value to the value created after splitting each branch (Lantz, 2013). These algorithms divide the sets into *k* subsets for *k*-fold cross validation and construction of a more refined model.

The nature of the approach to data splitting is dependent on the algorithm being employed. In this study, two different ML trees were used; the **Conditional Inference Tree (CTREE)** and the **Decision Trees using the C5.0 Algorithm (C5.0)**. For CTREE, the 'party' package was used (Horthorn et al., 2018), while for C5.0 the 'C50' package was used (Kuhn et al., 2018) and tuned in the 'e1071' package (Meyer et al., 2018).

Other decision trees exist that have proved successful in other studies, such as the AdaBoost Decision Tree for drug discovery (Korotcov et al., 2017) and the C4.5 algorithm, in modelling for cancer risk assessment (Quinlan, 1996; Richter and Khoshgoftaar, 2018).

Random Forests (RF). RFs are more robust than standard partitioning tree through the use of multiple tree constructions, finalising the model by finding the best fit through each group of trees (Breiman, 2001). The RF algorithm uses a small random partitioning of the data set to create each tree. This selection is then refined through multiple "out-of-bag" (OOB) operations, where the number of iterations try to minimize the error. Bootstrapping in this method is highly important for the forest to be formed out of as many robust trees as possible (Brownlee, 2016a). This particular technique provides a very powerful and easily trained classification model, reaching a similar complexity to SVM and NNET in many cases. DL approaches also exist to provide highly complex trees, however they are hard to interpret and may be complicated to construct.

For RF, the 'randomForest' package was used (Breiman and Cutler, 2018).

Partial Least Squares Discriminant Analysis (PLSDA). Traditional versions of this test are frequently used in GMM studies (Mitteroecker and Gunz, 2009; Mitteroecker and Bookstein, 2011), yet can be improved using ML algorithm training, such as the test presented here. This algorithm uses tuning of the internal algorithmic components to maximise accuracy, thus achieving in many cases higher accuracy than standard LDA methods (Mevik and Wehrens, 2007).

PLSDA, much like the *k* clustering model, works with Euclidian distances to calculate proximity with a group for classification. The maths behind the predictive modelling construction consists in multiple calculations. This algorithm then builds predictions via the centroids to each classifier and thus constructs continuously complex calculations in order to separate each group.

For PLSDA the 'pls' R package was used (Mevik et al., 2018).

Mixture Discriminant Analysis (MDA). MDA is similar to LDA, yet uses class-specific distributions to create a single multivariate classifier. The classes within the data set are divided for calculation of means and covariance structures, later prepared for a prediction model (Kuhn and Johnson, 2013). This model can be tuned using parameters according to the classes and subclasses that are divided and used for the training process.

The main difference with MDA is the means in which it treats the sample without assuming linearity, as opposed to LDA. Once the multiple class distributions are calculated, what MDA carries out is a mixture of the multiple distribution calculations to form a single classification component, mathematically described as:

$$D_n(x) \propto \sum_{k=1}^{L_n} \emptyset_{nk} D_{nk}(x)$$

In this equation, D_{nk} is the discriminant function for *k*th subclass in the *n*th class (Kuhn and Johnson, 2013:331). L_n Additionally is the number of distributions being used during the division and mixing part of the test (*ibid*).

For MDA, the 'mda' R package was used (Hastie and Tibshirani, 2017).

Naïve Bayes (NB). NB applies Bayes' theorem in the way it processes the estimation of class probabilities. Thomas Bayes, the mathematician who conceptualised this theorem, is one of the fundamental founders of multiple statistical and mathematical principles, especially in the field of probability calculation (Bishop, 2006). The NB algorithm uses training data to calculate the probability of each class based on the features presented by the input. Theoretically, this concept applies the conditional probability (*P*) that the final output is product of the variables (*A* and *B*) introduced, summarised as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}$$

The complexity of *A* and *B* however can vary greatly depending on the sample we have at hand (Lantz, 2013). Based off of the probabilities produced, the algorithm classes the input individual we wish to predict.

For NB the 'e1071' R package was used (Meyer et al., 2018).

K-Nearest Neighbour (KNN). Much like the *k* cluster method, KNN uses a very similar internal architecture to identify groups within the sample using primarily unsupervised pattern recognition of the sample. The use of KNN as a classification method is when the algorithm is provided with labelled sets, the model produced is closer to a reinforcement algorithm in nature, using the labelled cases to find similarities between the samples.

This classifier is a simple and sometimes effective algorithm, however, KNN is highly conditioned by the features presented and is susceptible to error when outliers and missing data occurs. KNN as a classifier as opposed to the unsupervised model presented before, transforms the data present to normalise the features present and induce a higher possibility for correct classification. The type of transformation used in this study is the z-score standardisation, explained mathematically as:

$$X_{new} = \frac{X - \mu}{\sigma} = \frac{X - Mean(X)}{StdDev(x)}$$

where X is the introduced dataset (Lantz, 2013).

KNN is considered a "lazy" algorithm because, unlike other ML techniques, the abstraction phase of learning is excluded (Kuhn and Johnson, 2013; Wei and Chiu, 2015). According to this technique, no learning actually occurs, meaning that no final model is created. KNN instead produces an algorithm that can classify based on the samples present, yet is highly reliant on the training set and requires "re-training" for every problem we wish to solve (Lantz, 2013).

For KNN, the 'class' R package was used (Ripley and Venables, 2015). The tuning of KNN is dependent on the *k* value presented and the standardisation procedure used. For this, the tuning parameters were defined using the 'e1071' package (Meyer et al., 2018).

These were the 9 models used specifically for their use as classifiers for GMM studies. Nevertheless, multiple other ML algorithms exist that can be used to train on data for both prediction and classification. Some of these were also included in other parts of this study and can include:

ML Regression Models. Standard regression models, as previously described, *can* be trained much like any other ML algorithm. The 'caret' package (Kuhn, 2018) provides a ML algorithm that trains a model based on regression data, providing the possibility of predicting values as well as classifying unknown data. A powerful extension of ML regressions are ML LDA, a refined version of LDA that is performed in the 'caret' package. In this version of LDA, the algorithm is trained based on traditional ML techniques for tuning and optimisation, using *k* fold cross validation and can use Bayes' theorem to optimise prediction probabilities.

Each of the models were prepared using a 70 : 30 % training/testing split. This sampling was carried out using the coded function presented by Wei and Chiu (2015), presented in Fig. 4.13.. All samples were bootstrapped 1000 times for training, ensuring that the small sample sizes (in some cases) could still be used for effective statistical analysis. This is a standard process to ensure efficiency of the final model. This was carried out using the 'boot' package in R (Canty and Ripley, 2017).

For ML based GMMs, the models were trained using the PC scores. The samples in the training process were resampled in order to estimate performance and adjust the weights of each model. This can be carried out using multiple methods, however in this case *k* fold cross validation was performed, with *k* set at 10. The tuning of each model was performed using

additional lines of R coding, carried out using the aforementioned packages for each test.

Models were evaluated considering 4 main statistics; Kappa, Sensitivity, Specificity and Balanced Accuracy. These are calculated using confusion matrices, performed through the `confusionMatrix()` function in the 'caret' package (Kuhn, 2018). The Kappa statistic adjusts accuracy through considering the possibility of a correct prediction by change along (Lantz, 2013). This is a common practice in calculus, and provides a value between -1 and 1, with any value above 0.8 considered as a powerful predictive model.

Sensitivity and Specificity tests combine different balanced evaluations of Type I and Type II errors. This ratio is a common practice in medical statistics (Fawcett, 2006). In order to calculate these values, the confusion matrix considers 4 fundamental concepts:

1. True Positives (*tp*)
2. True Negatives (*tn*)
3. False Positives (*fp*)
4. False Negatives (*fn*)

These 4 values are explained in Table 4.2.. Type I statistical errors are defined by the *fn* value, defining our ability to reject a hypothesis when it is in fact true, while Type II errors are the degree of *fp*s, whereby we accept a false null hypothesis. In medicine, sensitivity considers the likelihood of correctly diagnosing an illness, whereas specificity defines the likelihood of correctly identifying a healthy individual (Fawcett, 2006). These ratios are calculated as follows:

$$sensitivity = \frac{|tp|}{|tp| + |fn|}$$

$$specificity = \frac{|tp|}{|tp| + |fp|}$$

These calculations can essentially define the likeliness of misclassification (low value = high misclassification). Combining these two provides an additional calculation of balanced accuracy (Lantz, 2013).

Further means of evaluating performance, especially in the case of regression models, LDA and can also be applied to SVM, are the use of Receiver Operating Characteristic (ROC) graphs. ROC graphs have been a fundamental development in statistical development of predictive models, especially for fields of medicine, including investigation into HIV and other biological mutations (Beerenwinkel et al., 2002, 2003; Sing et al., 2005). The 'ROCR' package in R (Sing et al., 2005), provides a visual means of evaluating model performance via sensitivity/specificity results using *k* fold cross validation to produce a visual curve (Fig. 4.14.) that explains the relationship between the fp and tp rates. The greater area under the curve, the higher the model performance. This concept is called the Area Under the Curve value (AUC), with higher AUC values indicating better performing models.

An additional advantage presented by ML algorithms is the ability to evaluate model performance through an examination of variable significance. In the case of RF and Decision Trees, for example, this is considered a *pruning* of the predictive model. In cases where classification rates were low, this consequential pruning was carried out on the variables to try and see which variables could be used to optimise the model's classification abilities. In GMM, for example, common malpractice can be found in the assumption that the first PC or CV scores are the most important for sample differentiation. Theoretically, the first few PC scores represent the highest percentage of sample variance and covariance. However, analytical scores with lower percentiles can sometimes be the useful means of differentiating between groups, highlighting deviations from the majority (Albrecht, 1992). In cases where GMM PC scores presented a high number of scores representing low variability, the ML algorithms were tuned in accordance to the PC scores that were considered significant in training. This implies reducing statistical noise and thus avoiding overfitting.

Once models had been trained, tuned and tested, the best performing models were used for classification of the unknown values in FLK-West.

To understand the efficiency of these statistical computing methods, system performance was also monitored throughout the use of AI applications. This documented specifically the amount of computed processing power was required for the training of each model. An additional feature noted was the time lapsed when training algorithms, documented using the 'microbenchmark' package (Mersmann, 2018). Times were calculated over a total of 200 iterations per algorithm, while evaluation of processor performance was done over 30 iterations.

```
split.data = function(data,p = 0.7, s = 666) {
  set.seed(2)
  index = sample (1:dim(data)[1])
  train = data [index[1: floor(dim(data)[1]*p)],]
  test = data [index[((ceiling(dim(data)[1]*p)) + 1) :dim(data)[1]],]
  return(list(train = train, test = test))}

allset<-split.data(dataset, p = 0.7) # split 70% of data for training
trainset<-allset$train
testset<-allset$test
```

Figure 4.13. - R Code for creating a function that will split the dataset into 70% training and 30% testing sets for ML analysis

		Class	
		TRUE	FALSE
Prediction	TRUE	tp	fp
	FALSE	fn	tn

Table 4.2. - Table describing how a confusion matrix works, calculating the number of true and false positives as well as true and false negatives

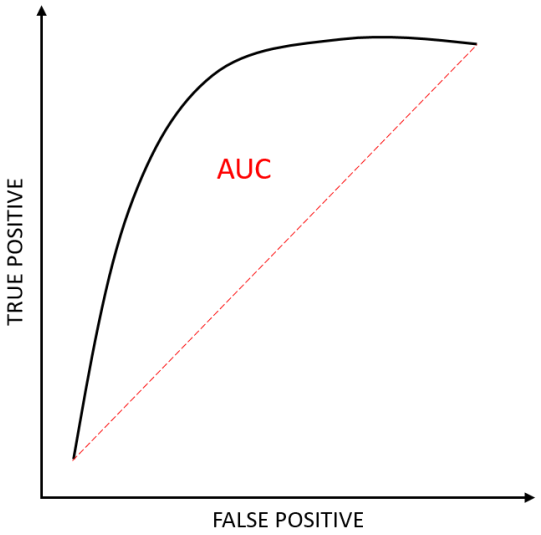


Figure 4.14. - Descriptive example of an ROC curve, with the AUC value indicated in red.

4.4.4. Deep Learning

One main type of DL algorithm was employed in this study; supervised **Deep Neural Networks (DNN)**. DNNs originate from standard NNET architectures and are currently the most widely used means of performing AI in multiple fields of research. DNNs originated primarily from their use in image-classification and have proven to be the most successful means of producing classification models. The main disadvantage of DNNs, however, are the complexity behind their original programming, requiring hours in some cases for their preparation before they can be used on any real data. Nevertheless, the success of DL over ML is seen in the much greater efficiency presented by DL algorithms once they are programmed. ML algorithms in many cases require the advanced preparation and refinement of data before their input into an algorithm. DL automates this task, learning the features presented and using multiple layers to fine tune the parameters required to best represent this information. This concept is known as *feature engineering* (Chollet, 2017).

For DNN, a separate digital work environment was prepared in Anaconda, isolating this environment specifically for DL tasks. Python was the sole application used for running DL models, preparing the DNNs internal architecture in VS Code for efficiency and debugging in a Jupyter Notebook. Both Conda (www.anaconda.com, Anaconda Inc, 2016) and PIP (http://www.python.org, Python Software Foundation, 2018) were used to manage the libraries and packages used within this environment. DNNs were constructed in Keras (Chollet, 2015, 2017), using a TensorFlow backend (Abadi et al., 2016). Keras is a library that provides the high-level building blocks for constructing deep learning models (Chollet, 2017; Chollet and Alaire, 2017). TensorFlow, on the other hand, is a powerful library that serves as a *backend engine* when performing DL tasks. A backend engines help the fluidity, movement and processing of data during DL applications. Multiple backends are available for DL, such as Theano (Theano Development Team, 2017, https://deeplearning.net/software/theano) and the Microsoft Cognitive Toolkit (github.com/Microsoft/CNTK), however for this study TensorFlow was employed (Abadi et al., 2016). TensorFlow has proven to be the most popular and powerful package developed by Google to manage and process the immense amount of data produced by this multinational company.

Four other Phython packages were used for DL, including Numpy (Oliphant, 2006; van der Walt, 2011), Pandas (McKinney, 2010), Scikit-learn (Pedregosa, 2011) and MatPlotLib (Hunter, 2007; Hunter et al., 2019). Numpy is an important component for scientific computations and the handling of large data arrays. Pandas can be used to import and transform this data for its use in DL, proving to be highly useful when dealing with different types of datasets and consequent classification problems. Scikit-learn is developed specifically for data mining and analysis. Finally MatPlotLib is a visualisation library designed for Python using Matlab interfaces, to produce an easy interface for graphic visualisation of these datasets.

Two different types of DNNs were designed, confronting two very different types of classification problems. The types of DNNs are directly dependent on the number of outputs required for the calculation of each model and can be defined as either binary or multiclass classification models.

Binary classifications were designed to respond to one simple question: is the taphonomic trace anthropic or not? In order to programme the network based on this question, the group labels were converted into binary code, with 1 being anthropic (cut or percussion marks) and 0 being non-anthropic (trampling or tooth marks). For multiclass classifications, this is slightly more complicated, requiring an entire

transformation of the output variables in order for the network to work. The most popular type of transformation in this case is called *one hot encoding*. This process reshapes the output attribute into a matrix with a Boolean data type according to each class value. This matrix transformation, much like binary, consists in a series of 0's and 1's. The encoder is then saved and used to decipher the output variables when making predictions. One hot encoding and decoding can be carried out using the `LabelEncoder` of Scikit-Learn (Pedrogosa et al., 2011).

The use of binary data rather than directly using categorical labels is quite simple to explain, whereby the network predicts the output based on the inputs and generates a single number that can then be rounded to produce the final classification response. The degree of rounding can then be used to infer the security of the prediction, and is later used for more complex evaluations of the models performance via a specified *loss* function.

As with ML, the inputs for these models were the PC scores extracted from GMM analysis, however in the case of DL a specific transformation of this data is required before it can be plugged in to any model. The input data is imported into Python as a list of values, which is then consequently converted into an array of numbers that can be indexed and sliced for multiple purposes. The first step consists in bootstrapping the samples to try and overcome any issues of sample size. The sample is thus bootstrapped 1000 times before being sliced to separate the input and output variables. The second step requires the splitting of these values into sets that can be used to train and evaluate the model.

As opposed to ML, which primarily uses Train and Test sets, DL splits the data one step further, creating a Validation set for model evaluation. The use of splitting the data into three separate sets in DL is due to the aforementioned *feature engineering* concept. In DL, the internal architecture of a model is composed of multiple components known as hyperparameters that can be tuned and refined to ensure the highest performance of the model. During the training and development of DNNs, for example, multiple phases consist in the configuration of these parameters and repeating the training step in order to tune the model based on a feedback signal produced from the validation data.

Every time a parameter is tuned, the model analyses its own performance based on the results used when analysing the validation set. During this retuning, information is leaked back into the model which creates a distinct phenomenon where the model "over learns" the data and begins to overfit. This process artificially produces good results, seeing how the model is optimised to the validation data, yet may produce completely different and unreliable results when used on unknown data (Chollet, 2017). In order to ensure this process does not happen, a third set is separated before being exposed to the model that can be used to evaluate model performance later. Multiple techniques can be used to perform model validation, including k-fold cross validation and tuning of internal parameters on a pre-split validation set before construction of the model (Brownlee, 2016b). In this study, both were trialled and tested, however for DL, the latter produced the best results and was used as the primary method in DNN evaluation.

Once the input and output data is prepared, the creation of a DNN consists in 5 main phases (Brownlee, 2016b):

1. Define network
2. Compile network
3. Fit network
4. Evaluate network
5. Predict using model.

DNNs in Keras are composed of a sequence of layers. The number of layers and nature of each is defined by the analyst and constructed in a "container" known as a *Sequential* class. In Keras, this simply requires constructing the model in an object class of this nature that can be compiled in the second stage of DNN construction. The only strict rules about layer types refers directly to the first and last layer of the network. This is simply

described by how the first defines the number of inputs while the last is used to mark the number of outputs (Brownlee, 2016b).

The hyperparameters presented in this stage of DNN construction refers directly to the model's internal architecture. The analyst is required to identify how many hidden layers are within the network, and the density of neurons contained within each of these layers. The transformation of data within these layers is then defined by the chosen activation function. While there is no strict rule of thumb for activation functions in most hidden layers, the final layer is normally what conditions the quality and type of final output. For binary classification, logistic activation or the 'sigmoid' function is preferred associated with 1 neuron that produces a class value between 0 and 1. For multiclass classification, the 'softmax' function is used in association with one output neuron per class value (Chollet, 2017; Brownlee, 2016b). For internal density layers within the network, the Rectified Linear Unit, or "relu", function is commonly employed.

These can mathematically be described as follows (Bishop, 2006; Heaton, 2012):

Relu: $f(x) = \max(0, x)$

Sigmoid: $f(x) = \frac{1}{1+e^{-x}}$

Softmax[1]: $f(x) = \frac{e^x}{\sum e^x}$

Two additional parameters that are usually effective for training are dropout and weight regularization techniques. A dropout layer randomly sets the output features to 0 during training, creating statistical noise in the process. While this seems counterproductive, this technique actually helps reduce overfitting by introducing noise, thus forcing the network to work harder in order to learn to overcome these issues and maintain the DNNs quality (Chollet, 2017). The weight regularisation component forces the network to only carry small values when refining the weights between neurons, thus making the network work harder to find a best means of representing the data without overfitting (Brownlee, 2019a).

The next step of the process consists in compiling the network. This converts the multilayer sequential model previously defined into a series of matrix transformations that compute the DNN. In order to compile the DNN, an additional series of hyperparameters are defined, in order to refine the training process that will be used during the next phases of construction. These parameters are usually tailored specifically to the classification task at hand. The two main hyperparameters that are defined in this stage are the *loss* function, the *optimization* function and the *metric* function. The *optimization* parameter defers to the way the model *learns* to fine tune the internal architecture. The *metrics* refers to the data collected when fitting the model. The *loss* function is a highly important component that is used to calculate the amount of misclassification produced during training, feeding this data back into the model to adjust the weights and fine tune the models performance based on the mistakes it makes.

The fitting phase of DL is probably the most important, and consists directly in the training process of the compiled model using all the previously defined hyperparameters to find the best means of representing the data. Training is performed using a *backpropagation* algorithm, whereby the metrics, optimization and loss functions are reintroduced into the system for fine tuning. The term backpropogation simply means that the error is passed back into the model, so that each algorithm can learn from its mistakes and thus adjust its weights. The *backpropagation* algorithm requires that the analyst define the data that the model will be trained on, the validation set and finally two additional

[1] For the purpose of this explanation, the mathematical formula for softmax has been simplified. The true formula can be found in Bishop (2006:198, 228)

hyperparameters known as the *epoch* and *batch* size. The *batch* is the number of samples that the DNN uses to train the model per *epoch*, while the *epoch* is the number of iterations performed during training. If we set the *epoch* to 100 and the batch to 32, then the DNN will repeat the training process 100 times with samples of 32 per iteration in order to tune the model. Too many epochs tends to produce overfitting, therefore the epoch number is generally modified on multiple occasions to find the best learning rate.

The final two steps of DNN construction simply consist in evaluating the networks performance, and once the optimum hyperparameter values have been discovered, used to finalise the model and use it to make predictions. The most popular evaluation method in DL consists in plotting the accuracy and loss during each epoch of training (Fig. 4.15.). This also helps detect over or underfitting, depending on the accuracy observed on validation samples which will then become increasingly apparent in the testing phase.

The standard methodology behind any DL study consists in experimentation with the different data sets, consequently adjusting all possible hyperparameters in order to find the best fit per model. In this study, different algorithms were tried and tested on models of varying layers and densities. Only one variable was changed per training session in order to see how each hyperparameter interacts with the models' performance. After a number of repetitions, the analyst usually gains enough experience to know which formulae work best in the case of each type of dataset. Most of the time this consists in trial and error, however, during the process of designing each DNN, an extensive bibliographical search consisted in consulting how other data scientists adapt to overfitting models and complicated data sets. For this, all the open-source platforms and libraries found online and in Anaconda proved highly important when designing the final DNNs presented here.

A total of 5 DNNs were constructed in this study. Some naturally performed better Consequent rejection as a solution to the equifinality problems present.

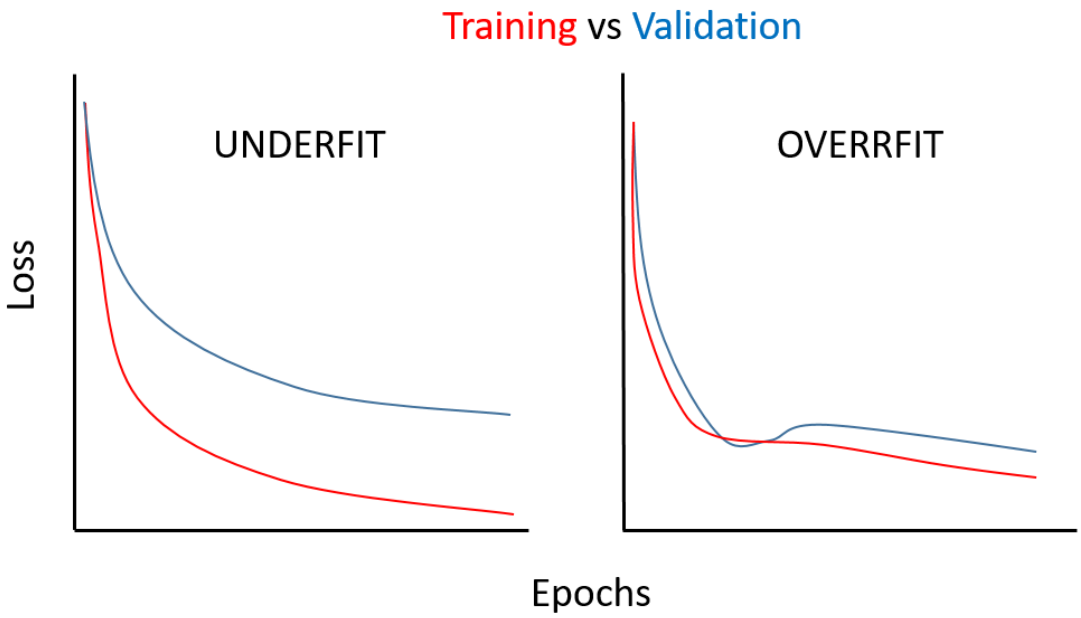


Figure 4.15. - Underfit vs Overfit loss curves plotted against number of epochs. Underfit models perform well on training sets yet are unable to classify efficiently the data from validation sets. Overfit models on the other hand show decent performance on validation sets to a point, and then a gradual decrease in performance. The perfect model would show both curves converging at the same point of the graph

Chapter 5

Results

5.1. Microscopy

The HIROX performed well, being able to successfully recreate each incision presented during the initial preliminary methodological study. The HIROX produces reconstructions in a time frame ranging from 5:32 to 20:50 min, averaging at 13:30 min (Appendix 2 – Table S1, Page 243). Compared with the time taken for the structured light scanner to process a mark, the reconstructions using the HIROX is a considerably longer process, however when compared with photogrammetry, the HIROX is still a more efficient means of producing high resolution images (Appendix 2 – Table S2, Page 244).

Comparing these times through multiple regression models, a step-wise regression was able to highlight a series of variables that condition the reconstruction time. These include, in order of importance, cut mark length, magnification, the number of tiles and finally the cut mark's width (AIC = -1.62). Comparison of AIC values and the use of multiple regression models was able to highlight that the impact all 4 variables have are significant ($p < 0.05$) on the processing time needed to digitally produce a 3D model of each mark. A clear correlation between these variables is easy to explain, considering how the number of tiles used is entirely dependent on the magnification, which in turn is conditioned by the length and width of the mark.

On the other hand, the time required for the landmarking process produced no correlation whatsoever, and can be seen to be product of experience rather than the technique being used.

Comparisons of lighting positions of the HIROX in contrast with the 3D reconstructions produced by the DAVID SLS-2 highlighted insignificant ($p \approx 1$) differences (Table 5.1.), proving that both techniques are statistically indistinguishable. 2D GMM analysis of cut mark cross sections produced PCA plots that overlapped greatly over all 10 PC scores (Fig. 5.1a.), whereas CVA results still present overlapping, despite the tendency this statistical test has in overestimating differences (Fig. 5.1b.). Nevertheless, in all cases, all 4 cut marks are easily distinguishable using all techniques (Table 5.2., $p < 0.05$).

When processing 3D data, a total of 19 PC scores were produced, generating a graph that shows how, regardless of the reconstruction technique, all 4 cut marks are easily distinguishable ($p < 0.05$, Fig. 5.2a., Table 5.2.). Furthermore, all 5 reconstruction techniques show significant ($p > 0.4$) similarities (Table 5.1.). These results are supported again in CVA, clearly being able to differentiate between marks while grouping all reconstruction techniques (Fig. 5.2b.).

Exploring variations in shape space for both reconstruction techniques, 2D analysis reveal how the lighting position mostly affects the depth of the mark (PC1), and the angle of the incision (PC2). This is logical to explain, considering the direction of the light would eventually affect the shadow being cast on different walls of the mark and the microscopes capacity to detect these features when creating the digital reconstruction. In the case of 3D analysis, these variations are considerably less noticeable, considering how this only truly affects the position of the 3rd landmark in Courtenay et al. (2017)'s model. The results observed through transformation grids are further supported by both numerical and graphical results, considering the best lighting position for digital reconstructions to be the lighting position from above. Furthermore, the mixture of coaxial and ring lighting were considered optimum for higher magnifications, therefore when processing superficial traces such as trampling marks, this technique was preferred for the rest of this study.

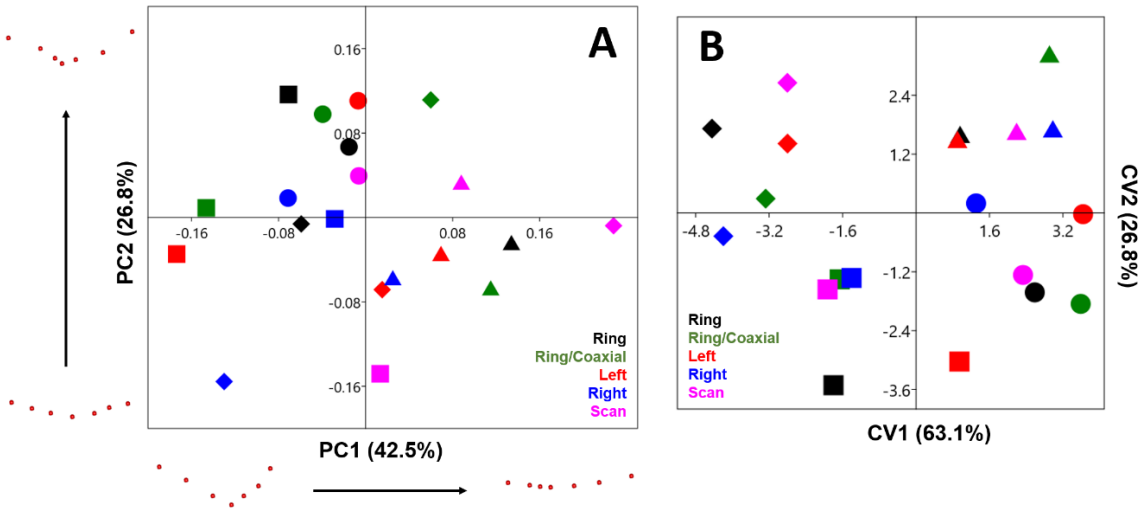


Figure 5.1. – Scatter plots presenting the (A) PCA and (B) CVA graphs comparing the 5 different reconstruction techniques for cut mark cross-sections. Each cut mark is presented as a different symbol. Variance in shape is presented for both extremities on the respected axis of each PC score.

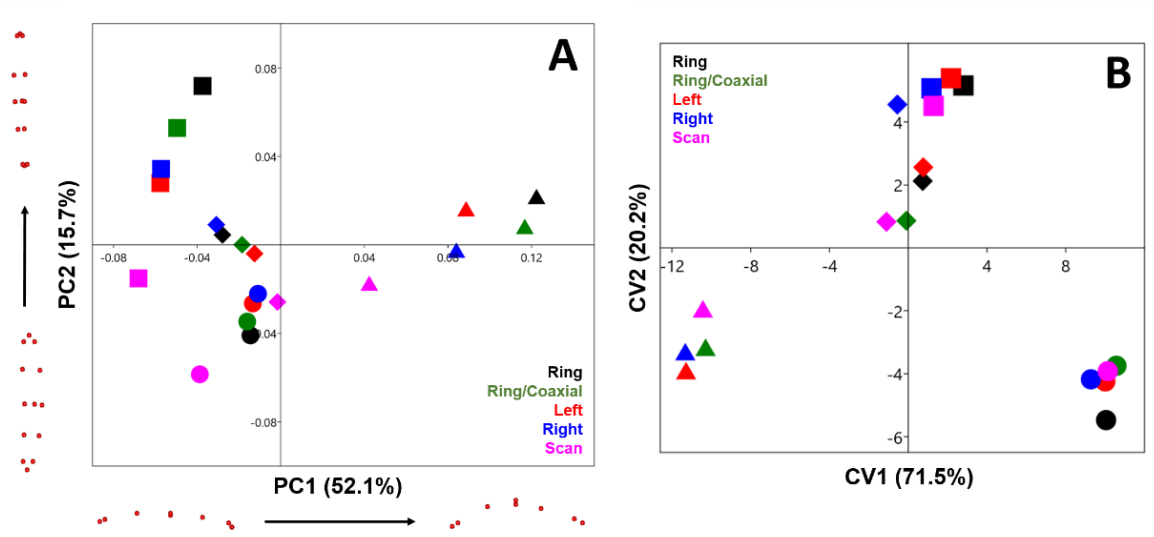


Figure 5.2. – Scatter plots presenting the (A) PCA and (B) CVA graphs comparing the 5 different reconstruction techniques for entire cut mark morphology. Each cut mark is presented as a different symbol. Variance in shape is presented for both extremities on the respected axis of each PC score.

	Ring	Mix	Left	Right	Scan
Ring		0.98689	0.86246	0.70995	0.85949
		0.94135	**0.77183**	**0.90411**	**0.40608**
Mix	0.98689		0.72821	0.62274	0.83721
	0.94135		**0.98715**	**0.99989**	**0.61758**
Left	0.86246	0.72821		0.9868	0.57261
	0.77183	**0.98715**		**0.99242**	**0.74994**
Right	0.70995	0.62274	0.9868		0.47208
	0.90411	**0.99989**	**0.99242**		**0.64374**
Scan	0.85949	0.83721	0.57261	0.47208	
	0.40608	**0.61758**	**0.74994**	**0.64374**	

Table 5.1. – MANOVA p values comparing the different reconstruction techniques, both using the 2D landmark model (in non-bold typeface) and the 3D landmark model (bold typeface)

	Cut Mark 1	Cut Mark 2	Cut Mark 3	Cut Mark 4
Cut Mark 1		0.6449	0.1569	0.1406
		0.0071	**0.0091**	**0.0015**
Cut Mark 2	0.6449		0.1567	0.1543
	0.0071		**0.0787**	**0.0026**
Cut Mark 3	0.1569	0.1567		0.4069
	0.0091	**0.0787**		**0.0075**
Cut Mark 4	0.1406	0.1543	0.4069	
	0.0015	**0.0026**	**0.0075**	

Table 5.2. – MANOVA p values comparing the different cut marks, both using the 2D landmark model (in non-bold typeface) and the 3D landmark model (bold typeface)

Once the statistical results had been analysed, the ideal conditions for 3D digital reconstructions was established, highlighting the following protocol for microscopic study of taphonomic traces:

1. After the main table is centralised and calibrated, the piece under study should be placed on a sturdy support with the area of interest positioned as flat as possible. This ensures that the piece does not move throughout the tiling process and prevents blurry photos.
2. The mark under study should be positioned as straight as possible along either a vertical or longitudinal axis, therefore reducing the number of tiles needed for reconstruction.
3. The mark under study should be positioned as flat as possible, thus reducing the focal depth required to capture the entire mark. While this is mostly conditioned by the natural topography of the piece, when considering high magnifications, this could greatly speed up the reconstruction process. The analyst is thus also advised to consult the depth of focus associated with each lens and adjust according to their needs.
4. In order to provide the highest quality 3D reconstruction, sufficient magnification is required to ensure the microscope is able to capture

the entire incision, especially focusing on the base, walls and surrounding cortical of the mark.

a. For 2D reconstructions, a magnification of approximately 400 to 600x is recommended (FOV = 505 µm) using the mid-range lens.
b. The cross section should be obtained between 30% and 70% of the mark's total length, as described by Maté-González et al. (2015).
c. For 3D reconstructions, the ideal magnification is normally considered between 100x (FOV = 1516 µm) and 200x (FOV = 3032 µm) magnification. This can be performed either using the low-range or the mid-range lens.

5. It is highly recommendable to ensure that a minimum of 30 photos are taken per tile, thus ensuring the best perception of depth for each mark.
6. The optimum lighting condition is positioned above the sample and combining both coaxial and ring light according to the magnification being used.
7. No polarised filters are advised at magnifications lower than 1000x

This protocol has recently been published as Courtenay et al. (2018b) in the Journal *Quaternary International*, available in Appendix 4.

5.2. Preliminary Experimental Trampling Study

The HIROX was highly successful when used to reconstruct trampling marks, however, during the process of analysing Yravedra (2005)'s 30 mark experimental trampling sample, the data began to produce interesting patterns that required further study. In continuation, these preliminary observations will be explained.

Through studying the 30 trampling mark sample by Yravedra (2005), 2 separate groups of trampling marks became gradually apparent based on a series of subjective qualitative criteria (Fig. 5.3.). Firstly, trampling marks seemed to vary in width, presenting a group of very wide marks in comparison with cases of very fine striae. Secondly, the number of internal striae within these marks also seemed to vary, with the wider marks presenting multiple internal striae along the base and walls of the incision while the thinner marks presented much fewer internal striations. In order to objectively characterise these marks, they were subject to a K-means clustering model (Table 5.3.), thus confirming that this preliminary sample contained two groups of very different trampling marks, with significant morphological differences (p = 9.12e-06).

PCA results and transformation grids were able to highlight how the true variables conditioning these cluster groupings lay directly in the width of the linear mark (Fig. 5.4.). Aditionally, the variance of shape and trajectory of each mark was able to confirm Domínguez-Rodrigo et al. (2009a)'s original characterisation of these linear incisions based on their sinuous or curvy groove trajectory.

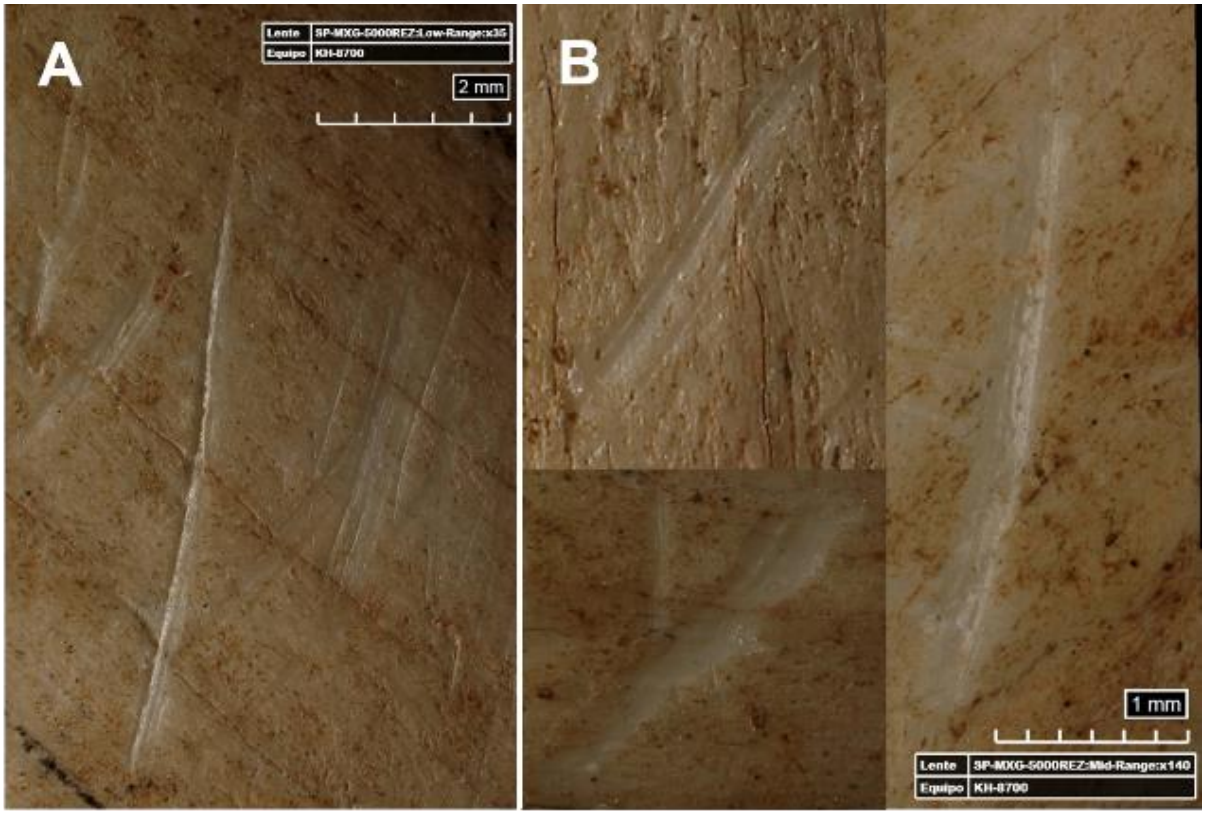


Figure 5.3. – Two different types of trampling marks identified preliminarily identified in Yravedra et al. (2005)'s sample. (A) Scratches and (B) Grazes

	Shape			Form		
K-Means	LDA	Graze	Scratch	LDA	Graze	Scratch
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	31%	69%	Graze	66%	34%
Scratch	Scratch	2%	98%	Scratch	27%	73%
Graze	Graze	83%	17%	Scratch	83%	17%
Scratch	Scratch	0%	100%	Graze	9%	91%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Graze	Graze	100%	0%	Scratch	20%	80%
Graze	Scratch	47%	53%	Graze	86%	14%
Scratch	Scratch	5%	95%	Scratch	1%	99%
Graze	Graze	97%	3%	Scratch	14%	86%
Scratch	Scratch	22%	78%	Graze	100%	0%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Graze	Graze	100%	0%	Graze	100%	0%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Graze	Graze	69%	31%	Graze	78%	22%
Scratch	Scratch	10%	90%	Scratch	7%	93%
Scratch	Scratch	2%	98%	Scratch	5%	95%
Scratch	Scratch	1%	99%	Scratch	5%	95%
Scratch	Scratch	1%	99%	Scratch	1%	99%
Scratch	Scratch	0%	100%	Scratch	6%	94%
Graze	Graze	99%	1%	Graze	94%	6%
Graze	Graze	98%	2%	Graze	87%	13%
Graze	Scratch	1%	99%	Scratch	29%	71%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	1%	99%	Scratch	10%	90%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	5%	95%	Scratch	11%	89%
Scratch	Scratch	0%	100%	Scratch	0%	100%

Table 5.3. – K-Means cluster groupings alongside LDA Classification/Misclassification values for each type of trampling mark.

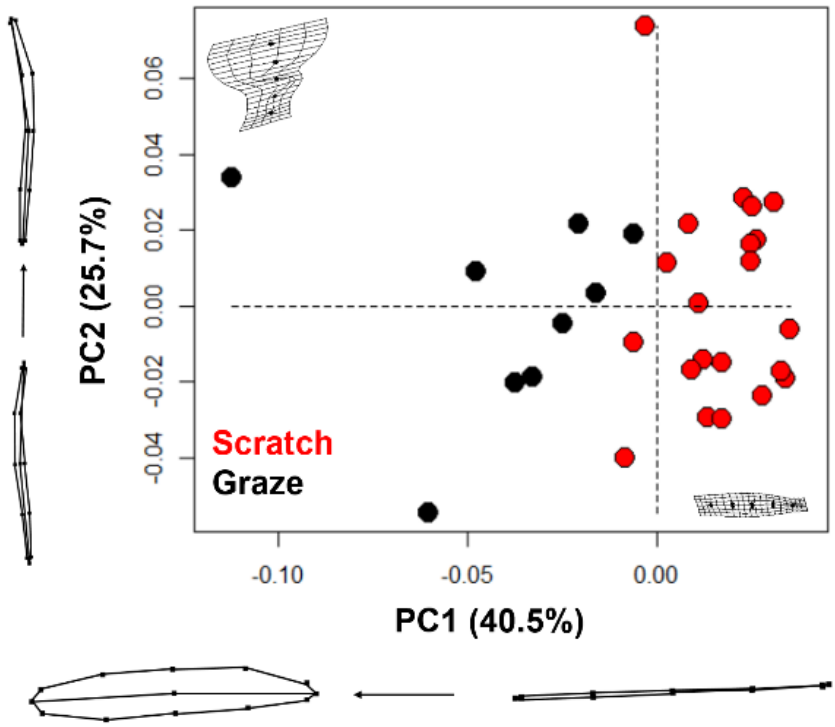


Figure 5.4. – Scatter Plot of the PCA results obtained when comparing trampling mark morphology in shape space. Variances in shape are represented on either extremity of their corresponding PC score axis.

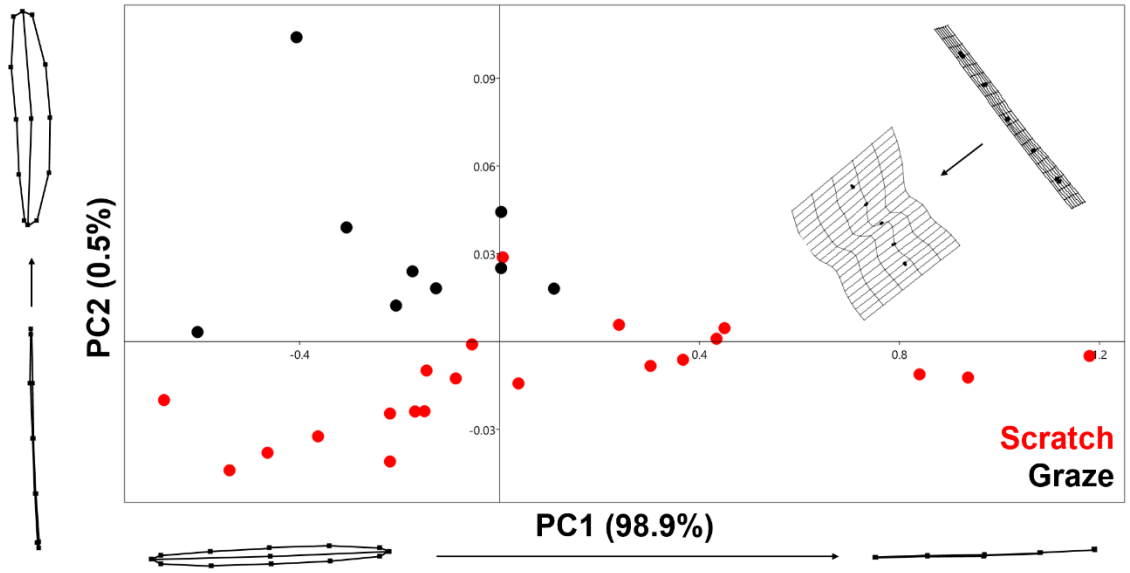


Figure 5.5. – Scatter Plot of the PCA results obtained when comparing trampling mark morphology in form space. Variances in form are represented on either extremity of their corresponding PC score axis.

CVA results were also able to clearly sepearte these groups, presenting significant Mahalanobis (D = 19.7491, p < 0.0001) and Procrustes (D = 0.0570, p = < 0.0001) distances. LDA results were also able to classify 93.33% of the different samples.

Form results also presented significant differences (p = 1.0543e-05) with a 76.67% classification between groups. Differences here were highlighted using transformation grids to be located in the distance between landmarks 10/11 and 12/13, possibly indicating a variation in length of the grooves (Fig. 5.5.).

The thinner of these incisions were consequently labelled "*scratch*" (n = 21) (Fig. 5.3a.), considering the etymological definition of this word as a "score or mark [of a] surface with a sharp pointed object (The Oxford English Dictionary, 2018). The second of these marks (n = 9) was provided the label "*graze*" (Fig. 5.3b.), to describe the "scrape or break [of a] surface lightly in passing" (The Oxford English Dictionary, 2018).

In order to ensure that this phenomenon was not exclusive only to Yravedra (2005)'s sample, the next phase of this study considered the analysis of a much larger experimental trampling sample, such as that of Domínguez-Rodrigo et al. (2009a). Once a sufficient sample size was collected, the true morphological characterisation and analysis of these marks analysed.

5.3. Trampling Study

For GMM, Domínguez-Rodrigo et al. (2009a)'s 251 trampling mark sample were subjected to a PCA which produced a total of 32 PC scores. The first two of these represent up to 58.4% of overall sample variance, while the first 10 represent up to 95% of total variance.

From this sample, a total of 131 scratches and 120 grazes were identified using an unsupervised K-Means clustering model on the PC scores produced during PCA. Analysis of K-Means clustering using ML based approaches highlighted how no significant change in grouping accuracy appeared after 2 iterations of the k-means loop (Fig. 5.6.), thus proving k = 2 to be the optimum model for pattern recognition in the analysis of trampling marks. This proves the aforementioned theory that two different types of trampling marks exist.

With this significant sample size, both types of trampling marks could be officially characterised using Geometric Morphometric methods. Additionally, an in depth study into the variables responsible for these morphological variations amongst trampling marks could be assessed.

GMM analysis of morphological variations through thin plate splines (Fig. 5.7.) revealed PC1 (33%) to describe the width of the mark while PC2 (25%) is characterised by the trajectory and curvature of the groove. Interestingly, PC's 3 through to 5 continue to represent the trajectory of the groove as the conditioning variable in separation, highly supporting the use of groove trajectory as a means of classifying trampling marks. Nevertheless, in all cases, transformation grids also support the superficial nature of these traces as a key component to their morphology.

MANOVA results highlight significant differences between groups (p = 0.001) while CVA revealed significant Mahalanobis (D = 3.79, p < 0.0001) as well as Procrustes (D = 0.0573, p < 0.0001) distances as well. Classic LDA approaches were able to successfully distinguish between groups with a 76% classification rate. While the larger sample size seems to have produced a larger overlapping of the samples and bring the two groups closer together, significant differences are still present between the two, highlighting their differences as taphonomic traces.

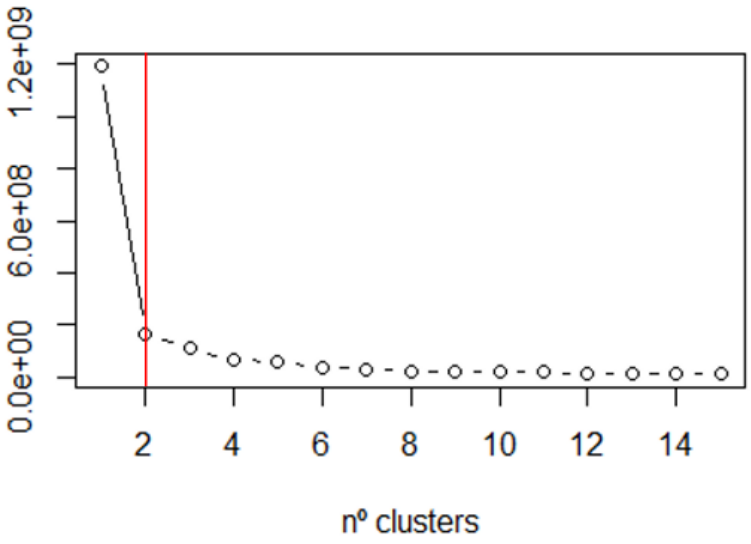


Figure 5.6. – Results of the unsupervised K-Means cluster algorithm looped evaluation using the Elbow-Joint method. The optimal clustering group size is k = 2 as indicated by the red line.

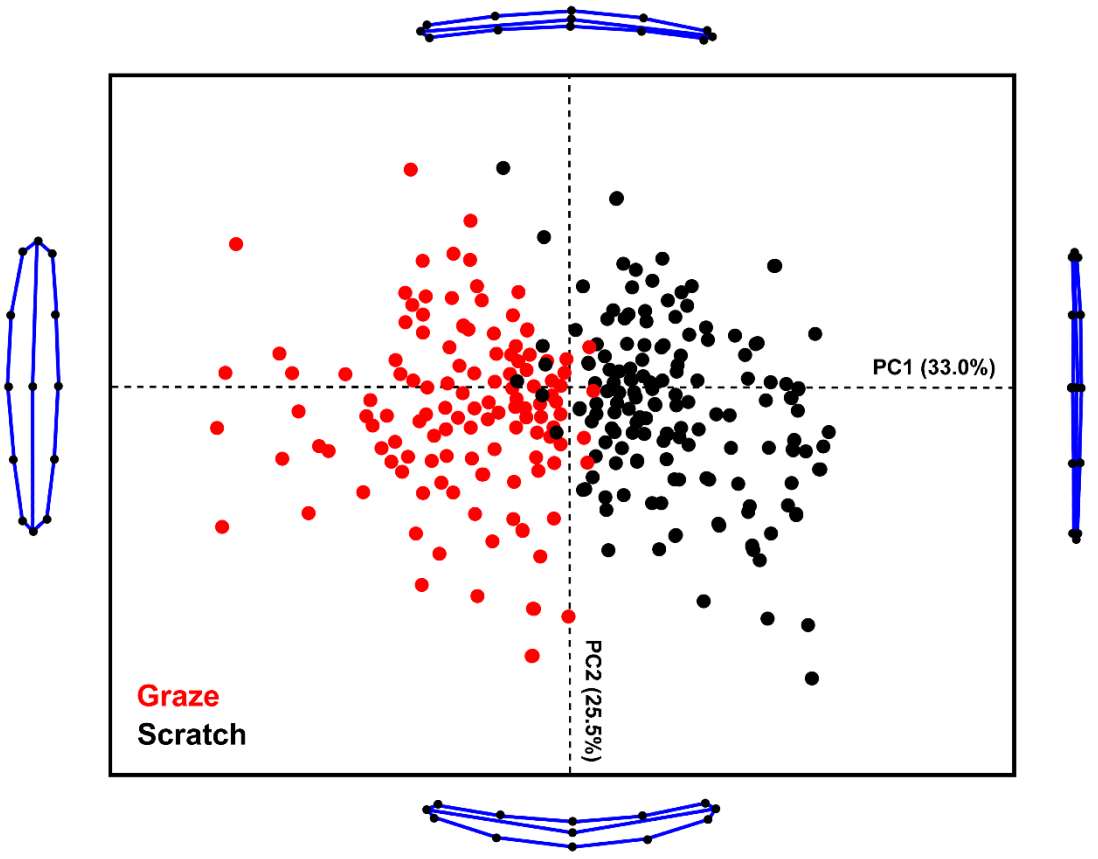


Figure 5.7. – Scatter plot presenting the PCA in shape space of the two different types of trampling marks. Variances in shape are represented on either extremity of their corresponding PC score axis

In form similar results can be observed with significant differences between groups (MANOVA, $p = 0.001$) and 89% differentiation between groups in LDA. Nevertheless, form space in this larger sample does not differ from the original study, thus little else can be extracted from this data. Much like in the preliminary study (Fig. 5.5.), grid warpings present a difference in relative position of landmarks 12/13 and 10/11, indicating a change in length of the marks as they become thinner (Fig 5.7.). Nevertheless, as noted by Courtenay et al. (2017), little can be said regarding form space when considering linear traces.

Turning to metric analyses of these traces, studies regarding the dimensions of these marks identified an abnormal distribution in both the length ($p = 7.47e-06$) and width ($p = 0.01$) of both types of incision (Table 5.4., Fig. 5.8.), additionally presenting a significant positive correlation for both scratches (rho = 0.35, p = 4.68e-05) and grazes (rho = 3.27, p = 0.00027). Plotting this data out shows a very clear pattern that separates the groups (Fig. 5.9.), producing general trend towards either short/wide marks or long/thin marks. 95% confidence intervals also highlight clear separations amongst groups (Table 5.4., Fig. 5.8.).

These finds objectively conclude that scratch marks are characterised by their long and thin nature as opposed to the short and wide graze marks.

The depth of the marks for both scratches (Shapiro-Wilk $p = 0.0001585$) and grazes (Shapiro Wilk $p = 0.000238$) are significantly different (ANOVA = 0.033), with no overlapping when considering 95% confidence intervals (Table 5.4., Fig. 5.8.). Through this it can be seen how grazes are deeper (31.96 +/- 4 µm) than scratches (19.70 +/- 2 µm).

Performing a TOST equivalency test on all three variables produced significant ($p = 0.00293$) results, highlighting that when combined, all three variables are not sufficient in classifying these taphonomic traces (Fig. 5.10.), however, comparing ANOVA tests on each variable in turn was able to distinguish how Depth in reality is the least effective means of testing for differences between samples, while width and length are much more significant (Table 5.5.).

If we are to remove the variable depth, calculations of differences are much more significant (MANOVA $p < 0.001$).

Multiple regressions on the variables width and length were performed to see the degree of classification produced by these two variables. The models produced, including linear regression (AIC = -64.9, AUC = 0.672), polynomial regression (AIC = -64.9, AUC = 0.672) and robust linear regression (AIC = -64.8, AUC = 0.684), were all able to produce powerful fits on the data (Fig. 5.11.), with sensitivity values of 0.85, specificity values of 0.86 and a balanced accuracy of 85.7%. Logistic regression however produced the best results (AIC = 60.6, AUC = 0.91), being able to provide a line of best fit that separated 98.5% of the samples (sens. = 0.97, spec. = 1).

It can clearly be seen across the four regressions how performance improved greatly when adjusting the mathematical function fitting the data, reaching a final model that represented almost the entirety of the data set (Fig. 5.11.). An in depth evaluation of the residuals produced by the final logistic regression proved that the model was able to fit the data almost perfectly to the prediction line (Appendix 2 – Fig. S1, Page 245), with the exception of two stray points that were consequently misclassified in confusion matrices.

	Sample	n	Mean	Confidence Intervals		SD	Min	Max
				95 Lower	95 Upper			
Length	Graze	120	3.76	3.60	3.93	1.12	1.70	6.18
	Scratch	131	5.35	5.02	5.68	2.32	2.44	10.76
Width	Graze	120	0.35	0.33	0.37	0.12	0.13	1.70
	Scratch	131	0.23	0.20	0.25	0.15	0.06	0.64
Depth	Graze	120	31.96	27.68	36.24	28.52	1	100
	Scratch	131	19.70	17.23	22.17	17.21	1	64

Table 5.4. – Descriptive statistics of the metric variables studied when comparing the different types of trampling marks.

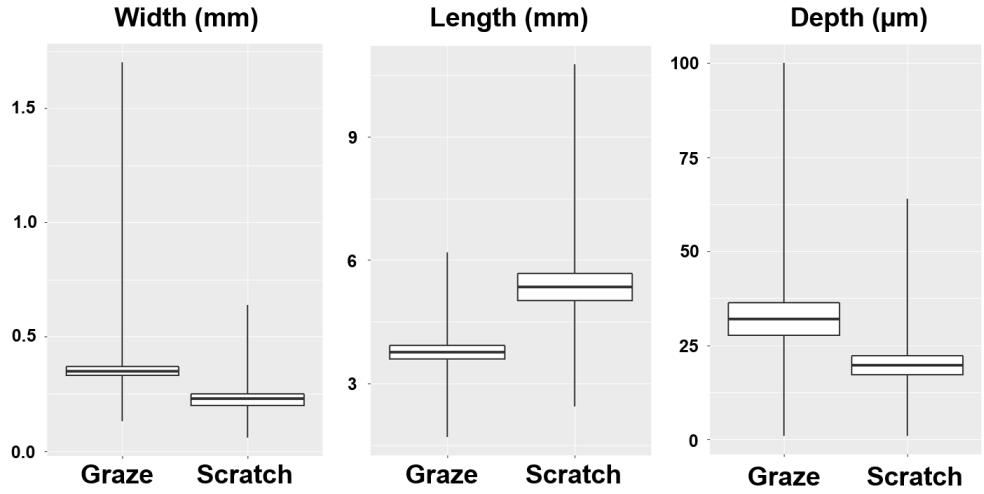


Figure 5.8. – Boxplots presenting the distribution of measurements when comparing the two different types of trampling marks. Width of the Box plots indicates

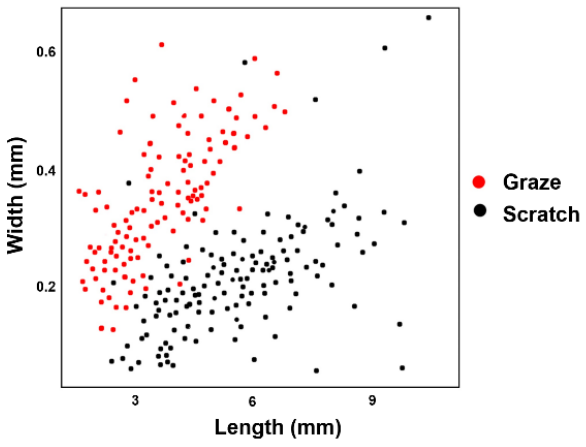


Figure 5.9. – Scatter plot of length and width measurements when comparing Graze and Scratch trampling marks

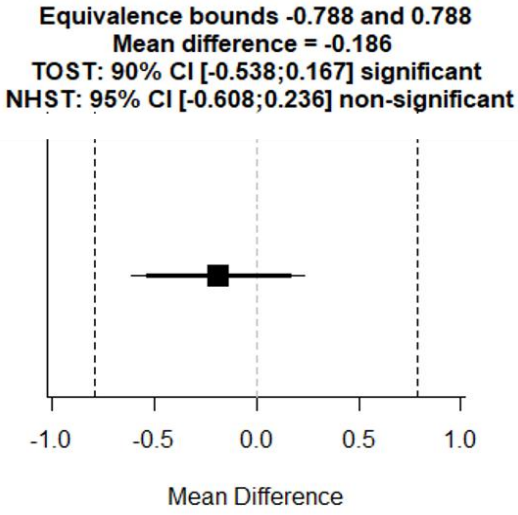


Figure 5.10. – TOST plot of mean differences and 95% confidence intervals present statistical equivalence between scratches and grazes when performing multivariate statistics on width, length and depth variables.

	F	*p*	Sig.
Width	16.37	0.000136	0
Length	11.84	0.000996	0
Depth	5.2355	0.02524	0.01

Table 5.5. – ANOVA results presenting the significance of the metric variables width, length and depth when comparing different types of trampling marks.

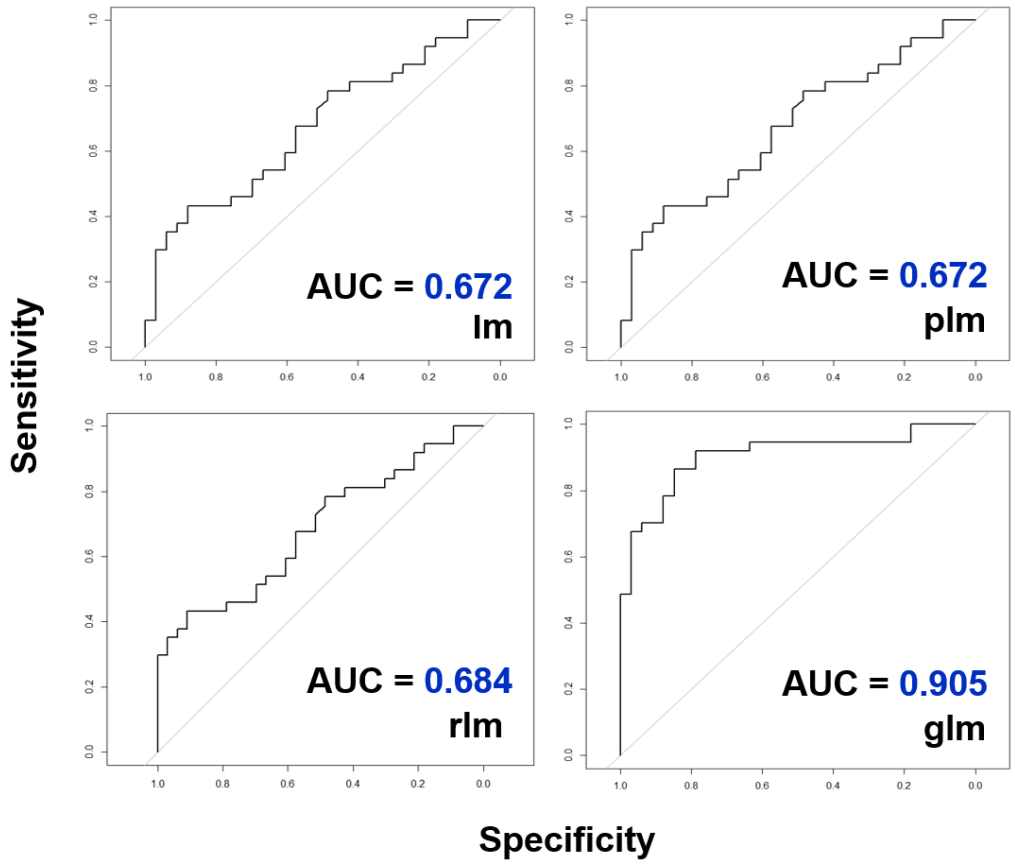


Figure 5.11. – ROC curves and AUC values for each type of regression performed on the variables length and width when differentiating between scratches and grazes. "lm" = Linear Regression. "plm" = Polynomial Linear Regression. "rlm" = Robust Linear Regression. "glm" = Logistic Regression.

Nevertheless, it can be seen here how scratches and grazes can be primarily characterised quantitatively through their width and length, helping in proving that a clear separation between these two cluster groups exists asides from the GMM perspective.

Using the variables proposed by Domínguez-Rodrigo et al. (2009a) a qualitative assessment was also performed on these marks. Of the 16 proposed variables, four main groups of observations can be considered important when analysing trampling marks. These include groove trajectory, mark orientation, overlapping striations and the presence and nature of internal micro striations.

The consequent results comparing the two different types of trampling marks highlight insignificant differences amongst most variables (Table 5.6.). This identifies the two different trampling samples to be practically identical qualitatively (Chi Squared test p = 0.5125). Two-sample t-tests between proportions were able to identify significant differences between a total of three variables;

presence of internal microstriations, trajectory of microstriations and the presence of microabrasions. Of these three variables, the presence of microstriations presented the highest significant differences between samples ($t(249) = 4.993$, $p = 0.0001$), followed by the presence of microabrasions ($t(249) = 3.519$, $p = 0.005$). Of these two variables, grazes were the sample to produce the highest quantity of both observations (microstriation n = 110 (91.8%), microabrasion n = 96 (80.0%)). The importance of the variable microstriation trajectory can be seen in a higher count of discontinuous microstriations in scratches (n = 32 (37.2%)), presenting significant differences when compared to grazes ($t(194) = 2.219$, $p = 0.0277$).

Nevertheless, the importance of these qualitative variables can arguably be criticised as highly subjective (Domínguez-Rodrigo et al., 2017b). This was confirmed in our study by significant differences ($p = 2.2e\text{-}16$) between the totality of our studied trampling sample and the original results produced by Domínguez-Rodrigo et al. (2009a) (Table 5.7.). The degree of subjectivity when comparing these two cases can be explained by our significantly higher perception of curvy trampling marks ($t(250) = 5.735$, $p < 0.0001$), as opposed to Domínguez-Rodrigo et al. (2009a)'s high count of sinuous marks ($t(250) = 6.282$, $p < 0.0001$). Significant differences have also been observed in qualitative descriptions of mark symmetry, presence of shoulder effect, presence of microabrasions, overlapping striae and the shape, location and nature of internal microstriations (summarised in Table 5.6.).

Considering how these variables are highly susceptive to subjectivity, the recording of associated microabrasions can thus be rejected as significantly important when identifying between trampling samples. On the other hand, it can be argued through significant similarities between both Domínguez-Rodrigo et al. (2009a)'s original observations and the results presented here that the variables concerning the presence of microstriations ($t(250) = 0.397$, $p = 0.6917$) and their continuity ($t(382) = 0.763$, $p = 0.4457$) are more reliable distinguishing factors. Regardless, recent studies have strongly suggested against the use of such subjective studies in qualitative features of marks, demanding for a more objective approach to the analysis of trampling marks, especially when compared with cut marks (Domínguez-Rodrigo et al., 2017b).

		Scratch	Graze	T-tests between proportions		
		n = 131	n = 120	T-Value	Freedomº	T-Probability
Groove Trajectory	*Straight*	47 (35.9)	41 (34.2)	0.282	249.0	0.7782
	Curvy	64 (46.6)	47 (39.2)	1.183	249.0	0.2381
	Sinuous	22 (16.8)	32 (26.7)	1.906	249.0	0.0578
Internal Microstriations	*Present*	86 (65.6)	110 (91.7)	4.993	249.0	<0.0001
	Absent	45 (34.4)	10 (8.3)	4.993	249.0	<0.0001
Trajectory of Microstriations	*Continuous*	53 (61.6)	86 (78.2)	0.2539	194.0	0.0119
	Discontinuous	32 (37.2)	25 (22.7)	2.219	194.0	0.0277
Shape of Trajectory	*Straight*	60 (69.8)	83 (75.5)	0.892	194.0	0.3735
	Irregular	25 (29.1)	28 (25.5)	0.563	194.0	0.5742
Location of Microstriations	*Walls*	9 (10.5)	19 (17.3)	1.349	194.0	0.1790
	Bottom	61 (70.9)	72 (65.5)	0.803	194.0	0.4227
	Both	15 (17.4)	20 (18.2)	0.145	194.0	0.8847
Mark Orientation	*Oblique*	109 (83.2)	102 (85.0)	0.389	249.0	0.6975
	Parallel	7 (5.3)	8 (6.7)	0.468	249.0	0.6404
	Perpendicular	14 (10.7)	10 (8.3)	0.646	249.0	0.5188
Microabrasion	*Present*	78 (59.5)	96 (80.0)	3.517	249.0	0.005
	Absent	53 (40.5)	24 (20.0)	3.517	249.0	0.005
Overlapping Striae	*Present*	65 (49.6)	60 (50.0)	0.063	249.0	0.9496
	Absent	66 (50.4)	60 (50.0)	0.063	249.0	0.9496

Table 5.6. – Count(Frequency%) table of the different qualitative features deemed to be important for characterising trampling marks according to Domínguez-Rodrigo et al. (2009a). T-test assessment between proportions are included to evaluate the significance of feature frequencies.

		This Study	DR (2009)	T-tests between proportions		
				T-Value	Freedomº	T-Probability
Groove Trajectory	*Straight*	88 (35.1)	75 (29.8)	1.045	250	0.2972
	Curvy	108 (43.0)	42 (16.7)	5.735	250	<0.0001
	Sinuous	54 (21.5)	134 (53.4)	6.282	250	<0.0001
Barb	*Present*	20 (8.0)	6 (2.4)	0.645	250	0.5196
	Absent	231 (92.0)	245 (97.6)	0.645	250	0.5196
Groove Shape	*\/*	36 (14.3)	10 (4)	1.214	250	0.2258
	_/	215 (85.7)	241 (96)	1.214	250	0.2258
Symmetry	*Symmetrical*	183 (72.9)	226 (90)	2.142	250	0.332
	Asymmetrical	68 (27.1)	25 (9.9)	2.142	250	0.332
Shoulder Effect	*Present*	71 (28.3)	15 (5.9)	2.799	250	0.0055
	Absent	180 (71.7)	236 (94.1)	2.799	250	0.0055
Flaking on Shoulder	*Present*	21 (8.4)	7 (2.7)	0.658	250	0.5114
	Absent	230 (91.6)	244 (97.3)	0.658	250	0.5114
Extent of Flaking	*Short*	16 (76.2)	2 (28.6)	2.278	26	0.0312
	Long	4 (19.0)	5 (71.4)	2.573	26	0.0161
Internal Microstriations	*Present*	196 (78.1)	188 (75.0)	0.397	250	0.6917
	Absent	55 (21.9)	63 (25.0)	0.718	250	0.4735
Trajectory of Microstriations	*Continuous*	139 (70.9)	169 (67.3)	0.763	382	0.4457
	Discontinuous	57 (29.1)	82 (37.2)	1.687	382	0.0925
Shape of Trajectory	*Straight*	143 (73.0)	140 (82.8)	2.31	382	0.0214
	Irregular	53 (27.0)	29 (17.2)	2.31	382	0.0214
Location of Microstriations	*Walls*	28 (14.3)	7 (2.9)	3.958	382	0.0001
	Bottom	133 (67.9)	219 (87.1)	4.49	382	<0.0001
	Both	35 (17.9)	25 (10)	2.228	382	0.0265
Mark Orientation	*Oblique*	211 (84.1)	206 (82.1)	0.523	382	0.6012
	Parallel	15 (6.0)	25 (8)	0.769	382	0.4424
	Perpendicular	24 (9.6)	20 (9.9)	0.099	382	0.9211
Microabrasion	*Present*	77 (30.7)	250 (99.6)	11.994	250	<0.0001
	Absent	174 (69.3)	1 (0.4)	23.153	250	<0.0001
Overlapping Striae	*Present*	125 (49.8)	203 (80.3)	4.397	250	<0.0001
	Absent	126 (50.2)	48 (19.7)	6.207	250	<0.0001

Table 5.7. – Count(Frequency%) table comparing the frequencies of observed qualitative features between this study and the study of Domínguez-Rodrigo et al. (2009a). T-test assessment between proportions are included to evaluate the significance of differences between the variables observed in this study and those observed in Domínguez-Rodrigo et al. (2009a).

Here the results regarding trampling marks are strongly supported by objectively quantitative data including geometric morphometric classification of each trampling mark into their corresponding group.

Once the sample had been fully classified, further investigation into the conditioning variables behind morphological variation was performed.

Each experimental variable considered in Domínguez-Rodrigo et al. (2009a)'s original study was tested for significance when considering the number of scratches and grazes identified. First of all, the weight of the person performing the trampling experiment was immediately rejected, producing no significant differences in the number of scratch or graze marks observed (*p* value per individual: SJ = 1.0, ABG = 0.2, MR = 0.2).

Neither did the variable time produce any significant patterns (t(250) = 0.934, p = 0.5265). This variable, however, did produce a significant correlation between the total number of trampling marks produced and the amount of time exposed to trampling (rho = 0.97, p = 0.033).

The different sedimentological conditions produced no significant changes to the number of scratch and graze marks produced (Table 5.8., Fig. 5.12.), however, did produce a very strong linear correlation (Fig. 5.13.) between the number of total trampling marks observed in relation to the size of the grain where it was buried (rho = 0.96, p = 1.06e-05). A general shift in total number of trampling marks can thus be seen to be directly affected by both the time exposed to trampling as well as the sedimentological conditions.

Considering the type of bone, no significance was seen in either axial (t(50) = 1.01, p = 0.32) or appendicular (t(35) = 0.361, p = 0.72) elements (Fig. 5.14.), however, highly significant differences were detected if considering the state of the bone when buried. Dry bones (t(101) = 3.897, p = 0.0002) can thus be seen to produce a much higher percentage of graze marks (n = 69 (68%)) while fresh bones (t(148) = 4.123, p = 0.0001) show a higher percentage of scratch marks (n = 98 (68%)). This almost perfectly inverse relationship (Fig. 5.15.) is thus the only variable that tests to be a significant (Table 5.8.) conditioning factor in the morphology of trampling marks.

Variable	Graze	Scratch	Graze %	Scratch %	T-Test Total		
					T-Value	Freedomº	T-Probability
Dry	69	33	68%	32%	3.897	101	**0.0002**
Fresh	51	98	34%	66%	4.123	148	**0.0001**
Fine-Grained Sand	7	4	64%	36%	0.967	10	0.3562
Medium-Grained Sand	10	9	53%	47%	0.262	18	0.7963
Coarse-Grained Sand	33	29	53%	47%	0.473	61	0.6377
Combination	24	31	44%	56%	0.896	54	0.374
Gravel	43	58	43%	57%	1.517	100	0.1324
Axial	29	22	57%	43%	1.01	50	0.3175
Appendicular	19	17	53%	47%	0.361	35	0.7205

Table 5.8. – Frequency of types of trampling marks considering the different variables analysed. T-Test assessment between proportions are included to evaluate the significance of these frequencies. The significant probability values are presented in bold type-face.

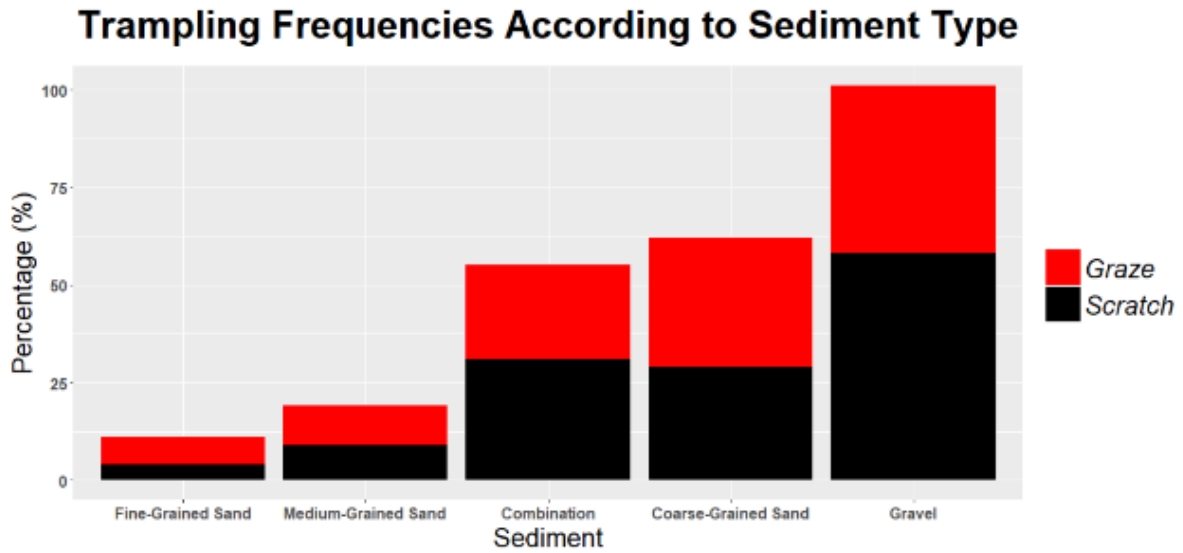


Figure 5.12. – Percentage of scratches and grazes according to different sedimentological contexts

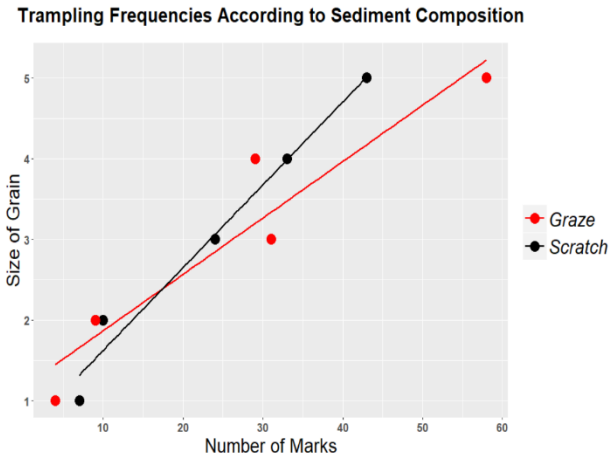


Figure 5.13. – Correlation of number of different types of trampling marks according to the size of grain of the sediment

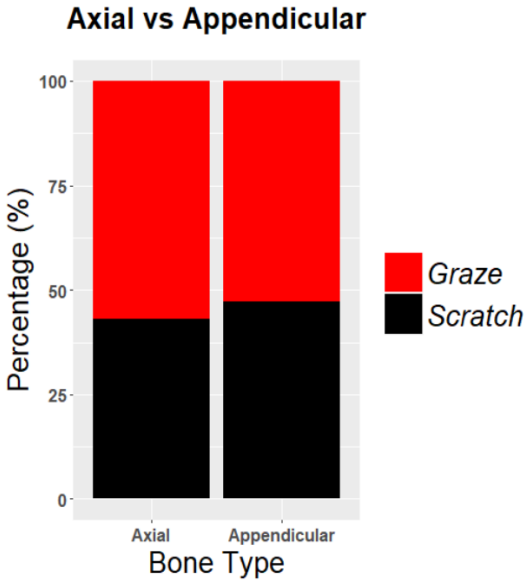


Figure 5.14. – Frequency of scratches and grazes according to the anatomical element

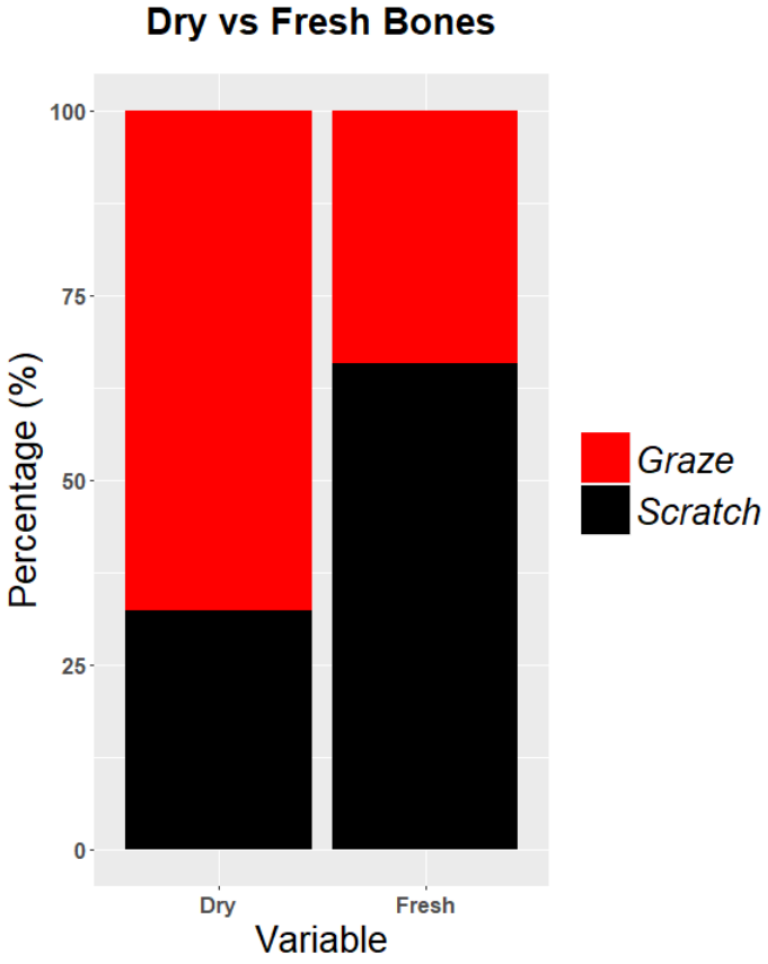


Figure 5.15. – Percentage of scratches and grazes according to the state of the bone (fresh or dry).

5.4. Trampling vs Cut Marks

The combined analysis of 80 cut marks with all 251 trampling marks produced a total of 32 PC scores in PCA. The first 2 PC scores represent a total cumulative proportion of 52% of the total sample variation and the following 8 represent up to 93%. When comparing the taphonomic traces without distinction between the different types of trampling marks, a great degree of overlapping can be seen over the first 2 PC scores (Fig. 5.16.). Nevertheless, a tendency can be seen for cut marks to be confined to a more constricted area of the overall feature space, while trampling marks present a great variability and appear spread out across the entire graph. This indicates that the overall variability of trampling marks is much greater. If we are to consider how this variability is distributed when separating trampling groups, the PCA begins to show much clearer patterns between the linear traces, with cut marks grouping more with scratches marks than with grazes. 95% confidence intervals drawn within this feature space (Fig. 5.17.) also shows a tendency for cut marks to fall closer to scratches, yet primarily occupy the first PC score, while the distribution across the second PC score is much more limited.

Exploring variations in shape across the PC scores shows how PC1 (31%) is represented through variations in mark width, while PC2 (21%) is represented through a variation in mark trajectory. It can be seen here how cut marks are much straighter and thus occupy a small percentage of PC2 while trampling marks vary in trajectory curvature and thus appear more spread out across both PCs 1 and 2. PC3 (17%) is interestingly represented through the depth of the mark (Fig. 5.18.), where cut marks begin to separate from the rest of the linear traces to the extremity of the PC score which is represented by deeper grooves as opposed to the shallower trampling marks.

Through calculations of mean shape using transformation grids (Fig. 5.19. and 5.20.), it can be seen how all linear traces are very similar in structure, highlighting the equifinality present when classifying these traces through simple observation. Grazes and scratches, when separated into their corresponding groups, show greater variation with cut marks than when lumped all together as trampling. Grazes as previously defined are much wider while scratches are much thinner. Cut marks fall in the middle of both, neither being too thin nor too thick, nevertheless, another significant difference can be seen in the depth of the incision (Fig. 5.19.), where cut marks are clearly characterised by a deeper groove.

MANOVA results are perfectly capable of differentiating between groups through significant differentiation of samples both when trampling marks are lumped together (p = 0.001) as well as when separated (p = 0.001 for all groups).

Mahalanobis and Procrustes distances between both groups are significant (Table 5.19.). CVA graphs (Fig. 5.21.) also highlight this difference with significant sample sizes that can be used to display the magnitude of variation and covariation. CVA results also highlight how the degree of variation in both trampling samples are much higher than in the cut mark sample, where cut marks occupy a much smaller area of the graph while both grazes and scratches are spread out across a large portion of the feature space. CVA performed on each group individually (Table 5.10.) also present a higher degree of overlapping between cut and scratch samples, however this is to a relatively minor degree (Appendix 2 - Fig. S2, Page 246)

LDA results produce confusion matrices with 81 +/- 8% classification between samples. When separating trampling groups, cut marks present a higher confusion rate with scratches (sensitivity = 0.73, accuracy = 78 +/-

10%) than they do grazes (sensitivity = 0.90, accuracy 84 +/- 9%). Graphical representation of these LDA differentiation can be found in Appendix 2 (Figures S3 and S4, Page 247).

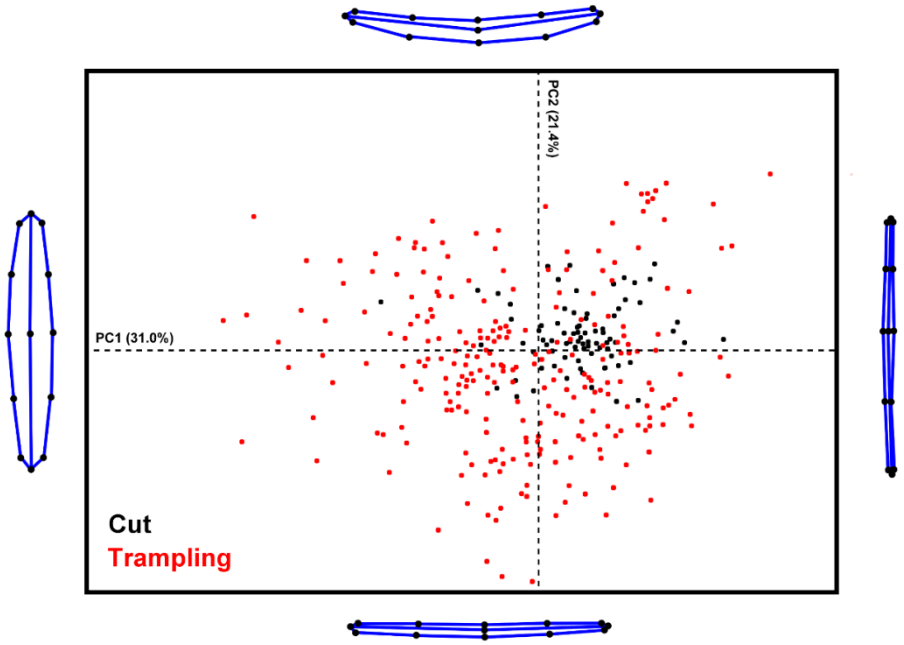


Figure 5.16. – PCA scatterplot presenting morphological variation between cut marks and trampling marks. Variances in shape are represented on either extremity of their corresponding PC score axis

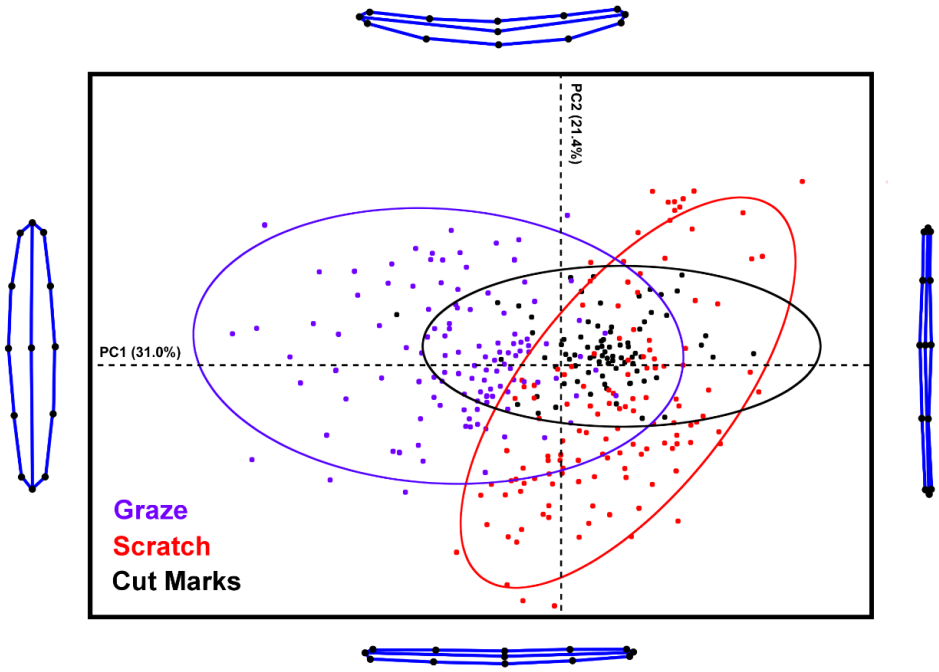


Figure 5.17. – PCA scatterplot with 95% confidence ellipses presenting morphological variation between cut marks and the two types of trampling marks. Variances in shape are represented on either extremity of their corresponding PC score axis.

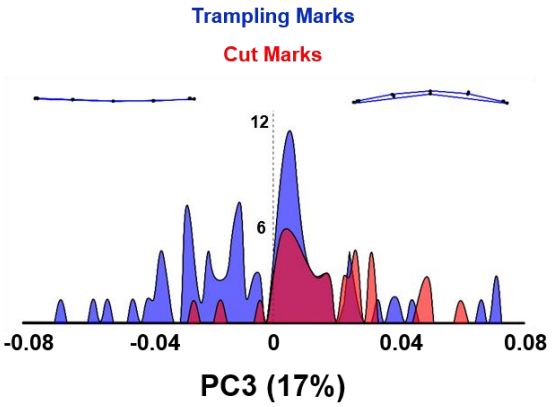


Figure 5.18. – Plot of frequency of points (y-axis) against the 3rd PC score, with shape variation represented on either extremity of the PC score.

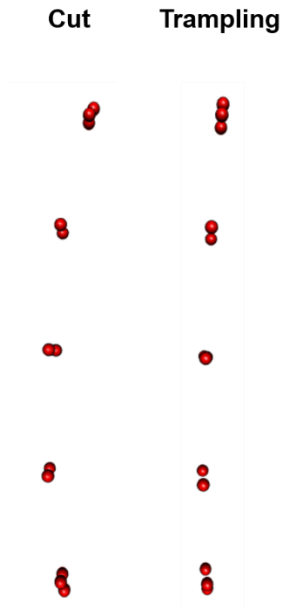


Figure 5.19. – Mean shape calculations for cut and trampling marks, representing a tendency for a deeper groove in cut marks as opposed to the shallower trampling marks.

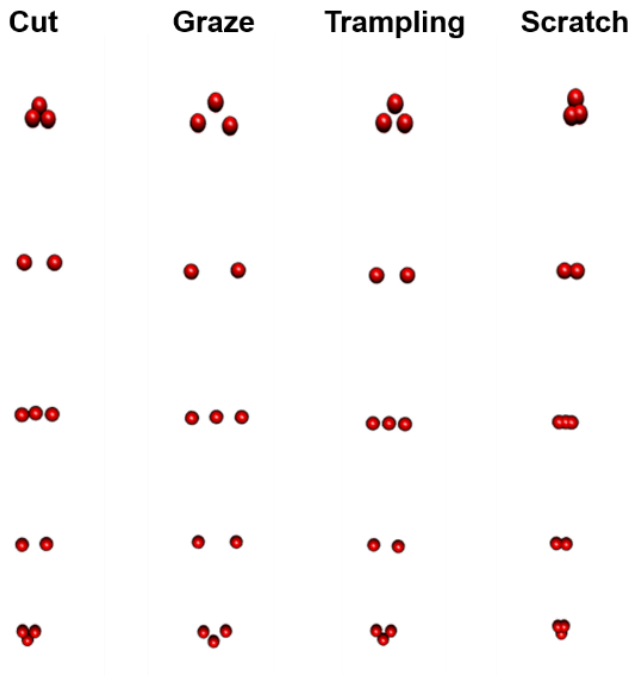


Figure 5.20. - Mean shape calculations for all linear traces.

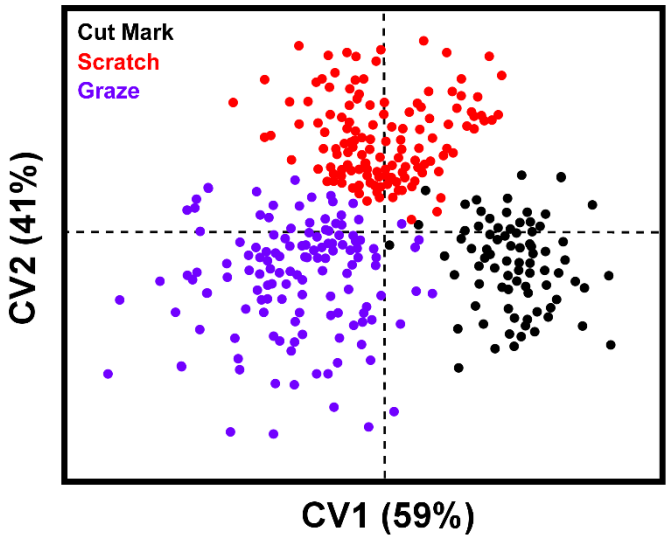


Figure 5.21. – CVA plot comparing the different types of linear traces.

	D	*p*
Mahalanobis	3.4403	< 0.0001
Procrustes	0.0297	0.0001

Table 5.9. –Mahablanobis and Procrustes distances between cut marks and trampling marks

		CM		Graze	
		Distance	p-value	Distance	p-value
Graze	Mahalanobis	3.9753	< 0.0001		
	Procrustes	0.048	< 0.0001		
Scratch	Mahalanobis	3.7127	< 0.0001	3.4637	< 0.0001
	Procrustes	0.0346	< 0.0001	0.0577	< 0.0001

Table 5.10. – Mahablanobis and Procrustes distances between cut marks and the different types of trampling marks

When comparing measurements taken for both scratches and grazes with cut mark samples (Table 5.11.), cut marks do not present any correlation between their length and width (rho = 0.216, *p* = 0.145) and present high degrees of overlapping with ANOVA results describing insignificant differences with scratches (length *p* = 0.07, width *p* = 0.58). With grazes samples present insignificant differences for length (*p* = 0.28), yet for width significant differences are observed (*p* = 3.673e-07) (Fig. 5.22.).

Nevertheless, when comparing the depth of trampling marks with cut marks produced with flint in natural butchery activities, significant differences in depth can be observed (ANOVA F = 15.968, *p* = 7.81e-07).

Reanalysing the metric variables that can be used to distinguish between trampling marks, TOST values still produce significant overlapping between trampling and cut marks (Fig. 5.23.), highlighting how metric simple measurements cannot be used to differentiate between the two.

While scratches and grazes can thus be best defined through their *length:width* ratio, this cannot be extended to include a differentiation of cut marks. Cut marks are best defined by their depth, yet as can be seen through these results, morphological comparison between the different traces can be considered more diagnostic.

			Confidence Interval					
		nº	Mean	95 Lower	95 Upper	SD	Min	Max
Cut Mark	Length	48	4.26	3.66	4.86	2.54	1.75	13.84
	Width	48	0.27	0.18	0.36	0.40	0.07	2.68
	Depth	48	62.00	50.78	73.22	47.26	7.00	222.00
Scratch	Length	131	5.35	5.02	5.68	2.32	2.44	10.76
	Width	131	0.23	0.20	0.25	0.15	0.06	0.64
	Depth	131	19.70	17.23	22.17	17.21	1	64
Graze	Length	120	3.76	3.60	3.93	1.12	1.70	6.18
	Width	120	0.35	0.33	0.37	0.12	0.13	1.70
	Depth	120	31.96	27.68	36.24	28.52	1	100

Table 5.11. – Descriptive statistics of the metric variables studied when comparing the different types of linear marks.

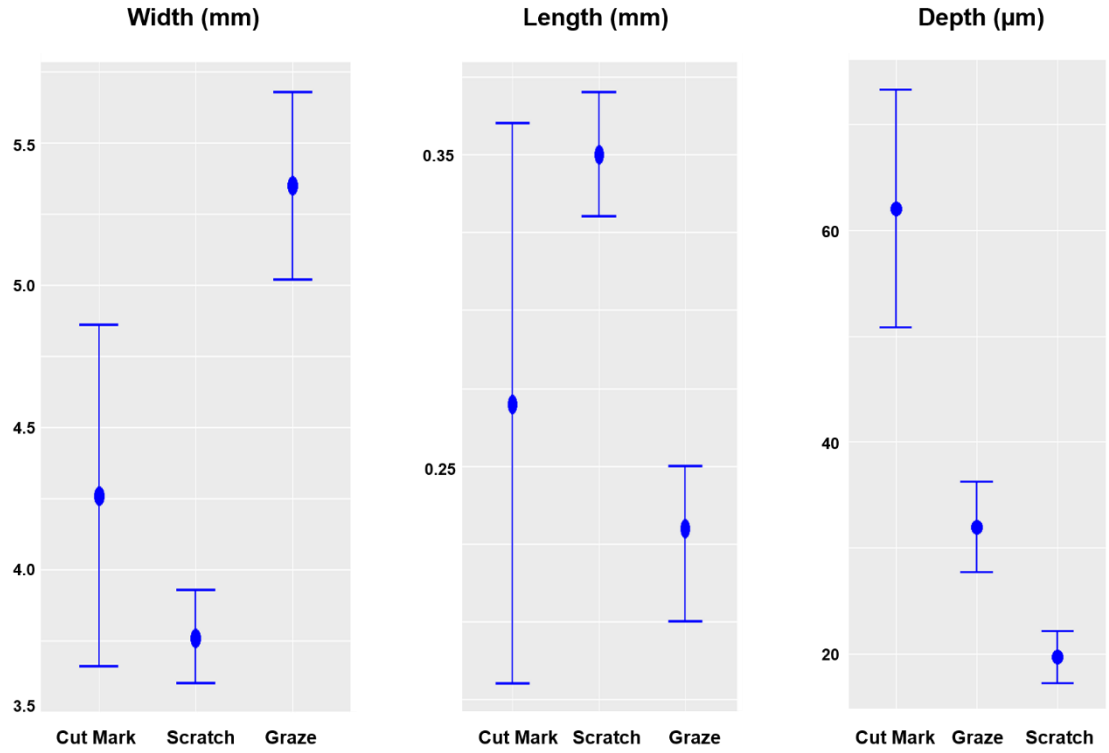


Figure 5.22. – 95% confidence intervals and mean (dot) of each metric variable for all three samples

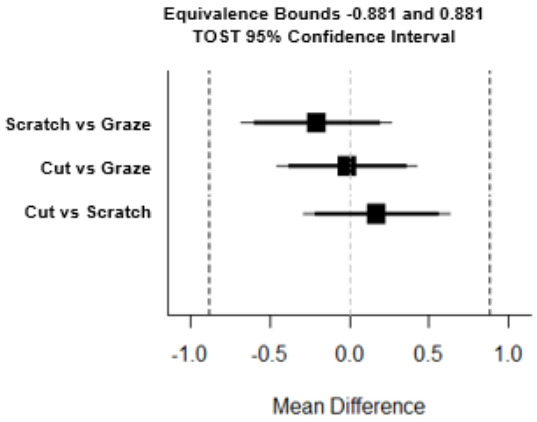


Figure 5.23. – TOST plot of mean differences and 95% confidence intervals present statistical equivalence between samples when performing multivariate statistics on width, length and depth variables.

5.5. Geometric Morphometric Characterisation of Different Carnivore Agencies

Seeing how GMM studies have already been described in detail by the original authors who published this data (Aramendi et al., 2017a; Yravedra et al., 2017c), this study only performed the initial steps of GMM in order to extract PCA data for the construction of classification models. Nevertheless, to provide a comparison with traditional GMM classification techniques with the new AI based models, LDA was also carried out.

PCA graphs comparing both carnivore tooth pits and scores present a very high degree of overlapping on most PC scores (Appendix 2 – Fig. S5 and S6, Page 248). For 3D analysis of tooth pits PCA produced a total of 44 PC scores in shape and 51 PC scores in form. In general, all PC scores represent a low degree of variance, with the exception of PC1 for form (77%). The remainder of PC scores throughout struggle to represent over 10% of total variance. For scores, the degree of variance represented by PCA is much greater, with as few as 10 PC scores for shape and 14 PC scores for form. The cumulative percentage of variance represented by the first two PC scores for both are as high as 89.3% for shape and 91.8% for form.

Calculations of mean shapes through thin plate splines present minute variations in morphology for each type of BSM (Fig. 5.24.), however slight variations can be seen regarding the depth and internal morphology of each tooth mark (Fig. 5.25.).

LDA for tooth scores show high confusion rates in shape space, with 49.6 +/- 8.99% accuracy, while in form space results improve but only very slightly, with an accuracy of 51.2 +/- 9.02%. In the case of tooth pits, confusion rates increase greatly with very low classification accuracy in shape (37.1 +/- 10.0%). Nevertheless, the best performance for distinguishing carnivore agencies is found in form (56.2 +/- 10.9%). Sensitivity and specificity values highlight a high degree of confusion in almost all cases, with the greatest degree of misclassification occurring for hyenas and wolves (Table 5.12.).

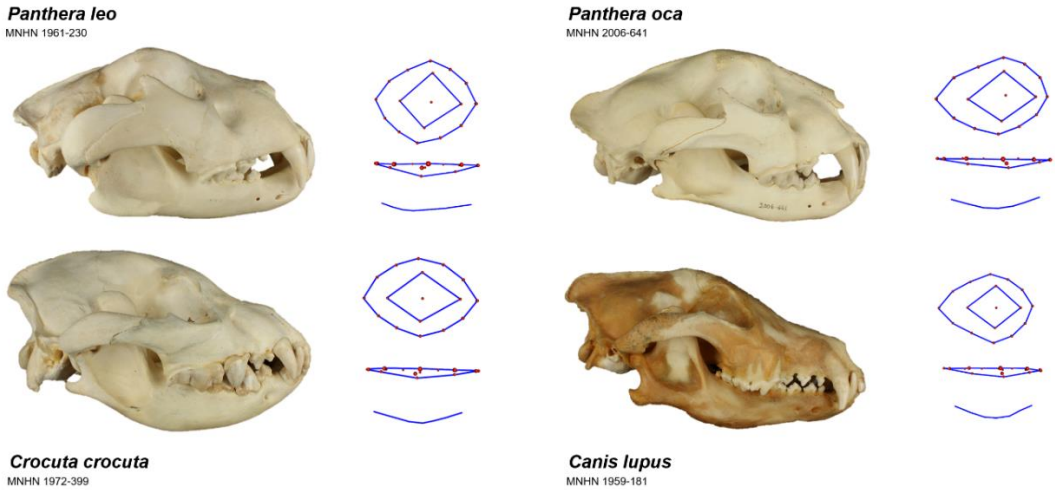


Figure 5.24. – Mean shapes of both pits and scores for each carnivore. All photos of skulls are from the Collections Mammifères et Oiseaux – Anatomie Comparée du Muséum National d'Histoire Naturelle, Paris (France).

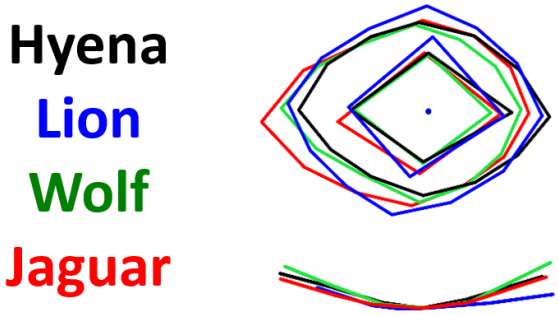


Figure 5.25. – Mean shapes calculated for both pits and scores. All shapes are superimposed using landmark 5 for pits and landmark 4 for scores.

			Hyena	Jaguar	Lion	Wolf
Scores	**Shape**	**Sensitivity**	0.39	0.56	0.63	0.40
		Specificifty	0.83	0.71	0.97	0.81
	Form	**Sensitivity**	0.30	0.62	0.70	0.43
		Specificifty	0.85	0.72	0.90	0.88
Pits	**Shape**	**Sensitivity**	0.57	0.20	0.33	0.38
		Specificifty	0.81	0.74	0.80	0.82
	Form	**Sensitivity**	0.62	0.50	0.58	0.54
		Specificifty	0.76	0.91	0.87	0.86

Table 5.12. – Sensitivity and Specificity values when comparing different carnivore tooth marks from LDA results.

5.6. Geometric Morphometric Characterisation of Percussion and Tooth Marks

For the differentiation of percussion marks and tooth pits, this study will describe the PCA data alongside LDA performance, as in the case of tooth marks described before, however for more detail regarding GMM characterisation of these traces, consult Yravedra et al. (2018).

PCA graphs (Appendix 2 – Fig. S7, Page 249) describing anthropic and carnivore activity present a greater degree of differentiation between tooth and percussion marks than in the case of carnivore taxa. Considering how the same GMM landmark configuration was used for tooth pits, the same number of PC scores were produced for both shape (n = 44) and form (n = 51) when comparing these two circular depressions. The overall degree of variance representation differs slightly (shape PC1 = 22.1%, PC2 = 16.3%; form PC1 = 77.1%, PC2 = 6.4%), however in general both shape and form are represented by a high number of PC scores with low percentages of represented variance.

Comparing mean shapes (Fig. 5.26.), tooth marks appear to be much more circular, with percussion marks more elongated. General differentiation between the two appears much clearer than in the case of different tooth marks produced by the multiple carnivores.

LDA performance on this data was much more efficient than for the different

carnivore tooth marks. General accuracy in shape produced 81.3 +/- 6.3% accuracy thresholds with decent specificity (0.85) and sensitivity (0.71) values. Form performed a little worse, with much higher degree of misclassification as seen through sensitivity (0.54) values. Nevertheless, specificity (0.90) is still high and the total model accuracy reached 78.9 +/- 6.7%.

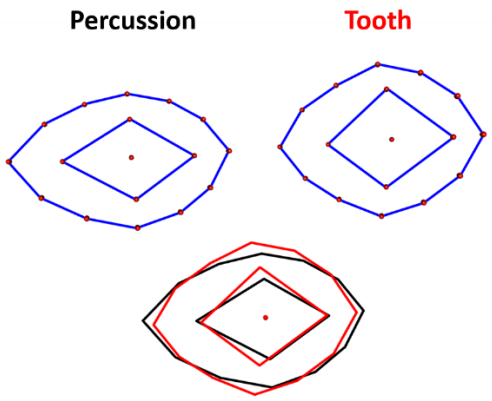


Figure 5.26. – Mean shapes calculated for percussion and tooth pits.

5.7. Machine Learning Classification Models

All supervised classification models were built from the top performing PC scores of each GMM model.

Considering the improvements in performance regarding shape and form for different models, these feature spaces were then used for model construction. Through this approach, PCA in shape space was employed to build classification models for linear marks and tooth scores, however form space was used to differentiate between carnivore taxa in tooth pits and percussion marks.

The ML models constructed to classify scratch and graze marks (Table 5.13.). Most models in this case were able to achieve 100% classification rates, with the best performing models including NNET, MDA, KNN, RF, C5.0 and SVM. The worst performing models were NB, CTREE and PLSDA, nevertheless, Kappa values highlight that these three algorithms are still powerful predictive models and still achieve over 90% accuracy in all cases. While the original LDA values presented an accuracy of 89% between the different types of trampling marks, Kappa (<0.8) values highlight how this particular model is not a strong classifier.

Model performance for the study of trampling marks proved successful in the cases of RF, C5.0 and SVM, all three producing significant Kappa and balanced accuracy values. Alongside this, 2 other models were able to produce powerful predictions (Kappa > 0.8), however accuracy values did not reach the 100% classification mark (acc = 0.97%). The remainder of models produced poor classifiers (Kappa < 0.8) with high levels of misclassification. While traditional GMM LDA results proved to distinguish a high percentage of cut and trampling marks (acc = 80%), it is important to consider the degree of misclassification, consequently producing an unsatisfactory Kappa value (0.5).

Through these two case studies, it can be seen how SVM, RF and C5.0 are the most powerful models for processing data of this type. LDA may produce high classification rates, yet great care should be taken, regarding how the Kappa statistic does not reach a value that can be considered satisfactory. Interestingly, NNET performance falls when greater degrees of overlapping are present, this quite possibly reflects on the need for a more complex internal structure to process highly overlapping samples. Other models, however, are relatively inconsistent, some producing high classification rates with powerful Kappa values, yet others fall below an acceptable threshold.

When constructing ML classification models for the processing of 2D 7-landmark

GMM data, RF, C5.0 and SVM continue to appear as the most powerful models (Table 5.15.). Nevertheless, many other algorithms are able to produce powerful predictive models. Interestingly, shape space appears to produce a higher number of models with lower misclassification rates than form space (Fig. 5.27.). Additionally, considering reflections made by multiple authors, the removal of wolves from the mix reduces statistical noise and produces higher degrees of classification (Fig. 5.28.).

When considering LDA results, on all accounts this model produces very poor classification results and should thus be discarded as a means of differentiating between carnivore agents via tooth scores.

For the case of 17-landmark 3D GMM models, similar results can be seen (Table 5.16.), yet with a higher degree of performance when considering form space (Fig. 5.27.). Considering the large number of PC scores, variable importance was evaluated and the models were tuned according to the top performing scores (Appendix 2 – Figure S8 and Tables S3 to S5, Page 249-251). It can be seen how the use of the top 15 PC scores produced the highest rate of classification, yet the presence of wolves still produce a significant degree of statistical noise. If wolves are excluded from analysis, model performance increases greatly (Fig. 5.28.).

Once the noise produced by wolves had been observed for wolves, it was decided that this sample be removed when classifying archaeological data in this particular study. To stay true to the carnivores present in the African savannah, the final models were tuned only on lion and hyena tooth marks.

Finally, when using this 17-landmark model to distinguish between anthropic and carnivore agencies via circular BSMs, NNET begins to appear as a powerful classifier alongside RF, C5.0 and SVM that continue to present the highest classification performance. Nevertheless, almost all models show very high kappa values that prove to be efficient classifiers (Table 5.17.).

	Performance		95% Confidence Intervals			
	Kappa	**Accuracy**	**Lower**	**Upper**	**Sensitivity**	**Specificity**
GMM LDA	0.77	0.89	0.80	0.95	0.79	0.97
NNET	1	1	0.99	1	1	1
NB	0.80	0.90	0.86	0.93	0.98	0.81
CTREE	0.96	0.98	0.96	0.99	0.97	0.99
PLSDA	0.88	0.94	0.91	0.97	0.88	1.00
MDA	1	1	0.99	1	1	1
KNN	1	1	0.99	1	1	1
RF	1	1	0.99	1	1	1
C5.0	1	1	0.99	1	1	1
SVM	1	1	0.99	1	1	1

Table 5.13. – Evaluation of different classification model performance distinguishing between scratches and grazes.

	Performance		95% Confidence Intervals			
	Kappa	Accuracy	Lower	Upper	Sensitivity	Specificity
GMM LDA	0.50	0.80	0.71	0.87	0.86	0.93
NNET	1	1	0.82	1	1	1
NB	0.60	0.80	0.75	0.84	0.93	0.67
CTREE	0.93	0.97	0.94	0.98	1.00	0.93
PLSDA	0.56	0.78	0.73	0.83	0.76	0.80
MDA	1	1	0.85	1	1	1
KNN	1	1	0.95	1	1	1
RF	1	1	0.99	1	1	1
C5.0	1	1	0.99	1	1	1
SVM	1	1	0.99	1	1	1

Table 5.14. – Evaluation of different classification model performance distinguishing between Cut Marks and Trampling Marks.

		Performance		95% Confidence Intervals			
		Kappa	Accuracy	Lower	Upper	Sensitivity	Specificity
GMM LDA	**Shape**	0.32	0.50	0.41	0.59	0.60	0.83
	Form	0.35	0.51	0.42	0.60	0.51	0.84
NNET	**Shape**	0	1	0.54	1	1	1
	Form	1	1	0.82	1	1	1
NB	**Shape**	0.82	0.87	0.82	0.90	0.89	0.96
	Form	0.64	0.73	0.67	0.78	0.73	0.91
CTREE	**Shape**	0.89	0.92	0.88	0.95	0.92	0.97
	Form	0.95	0.96	0.93	0.98	0.96	0.99
PLSDA	**Shape**	0.70	0.77	0.72	0.82	0.78	0.93
	Form	0.99	0.99	0.97	1.00	0.99	1.00
MDA	**Shape**	0.92	0.94	0.91	0.97	0.94	0.98
	Form	0.99	0.99	0.98	1.00	1.00	1.00
KNN	**Shape**	0.93	0.95	0.91	0.97	0.95	0.98
	Form	0.98	0.99	0.97	1.00	0.99	1.00
RF	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
C5.0	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
SVM	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1

Table 5.15. – Evaluation of different classification model performance distinguishing between different carnivores through tooth score cross sections. All sensitivity and specificity values are averaged between wolves, jaguars, hyenas and lions.

		Performance		95% Confidence Intervals			
		Kappa	Accuracy	Lower	Upper	Sensitivity	Specificity
GMM LDA	**Shape**	0.13	0.35	0.25	0.46	0.34	0.78
	Form	0.16	0.57	0.27	0.48	0.37	0.80
NNET	**Shape**	1	1	0.58	1	1	1
	Form	0	1	0.54	1	1	1
NB	**Shape**	0.47	0.62	0.58	0.69	0.60	0.87
	Form	0.66	0.75	0.69	0.79	0.75	0.92
CTREE	**Shape**	0.80	0.85	0.81	0.89	0.83	0.95
	Form	0.78	0.84	0.79	0.88	0.84	0.95
PLSDA	**Shape**	0.45	0.60	0.54	0.65	0.57	0.86
	Form	0.65	0.74	0.69	0.79	0.74	0.92
MDA	**Shape**	0.72	0.79	0.74	0.83	0.78	0.93
	Form	0.75	0.81	0.76	0.86	0.82	0.94
KNN	**Shape**	0.95	0.96	0.94	0.98	0.97	1.00
	Form	0.96	0.97	0.95	0.99	0.98	0.99
RF	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
C5.0	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
SVM	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1

Table 5.16. – Evaluation of different classification model performance distinguishing between different carnivores through tooth pits. All sensitivity and specificity values are averaged between wolves, jaguars, hyenas and lions.

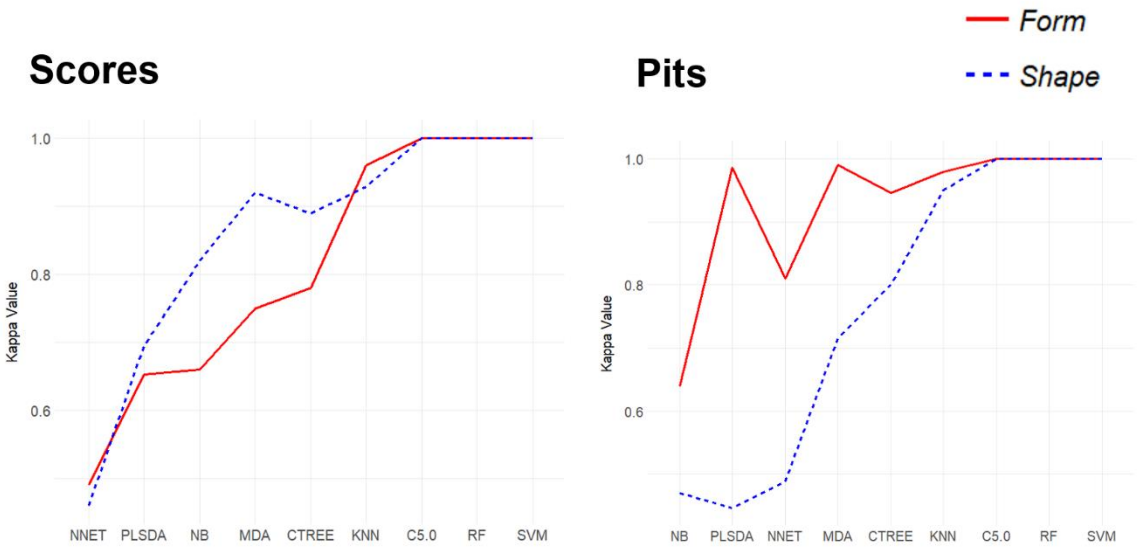


Figure 5.27. – Kappa value curves evaluating ML model performance on PC data in both form and space for different types of carnivore tooth marks.

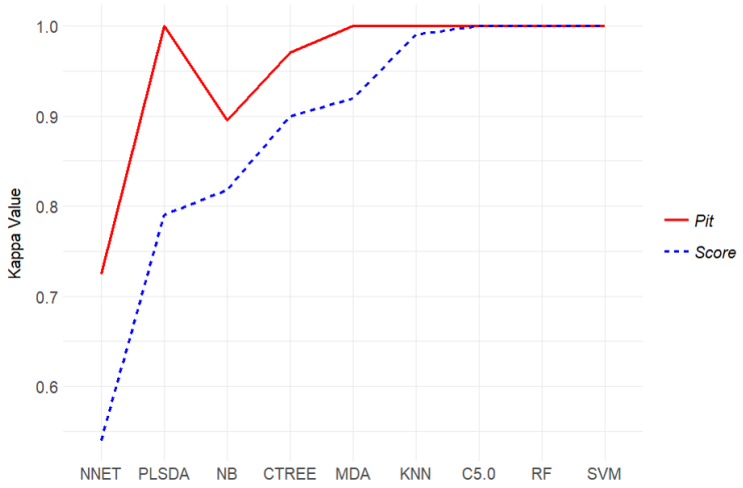


Figure 5.28. – Kappa value curves evaluating ML model performance on PC data once having removed wolves from the analysis. PCA data used to train models are in form for pits and space for scores.

		Performance		95% Confidence Intervals			
		Kappa	**Accuracy**	**Lower**	**Upper**	**Sensitivity**	**Specificity**
GMM LDA	**Shape**	0.56	0.81	0.73	0.88	0.72	0.85
	Form	0.47	0.79	0.71	0.86	0.54	0.90
NNET	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
NB	**Shape**	0.67	0.86	0.82	0.90	0.66	0.96
	Form	0.68	0.86	0.82	0.90	0.78	0.90
CTREE	**Shape**	0.90	0.95	0.92	0.97	0.90	0.98
	Form	0.88	0.95	0.91	0.97	0.92	0.96
PLSDA	**Shape**	0.79	0.91	0.87	0.94	0.79	0.97
	Form	0.96	0.98	0.96	0.99	0.95	1.00
MDA	**Shape**	0.94	0.97	0.95	0.99	0.95	0.98
	Form	0.95	0.98	0.95	0.99	0.93	1.00
KNN	**Shape**	0.96	0.98	0.96	0.99	1.00	0.97
	Form	0.90	0.96	0.93	0.98	0.99	0.94
RF	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
C5.0	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1
SVM	**Shape**	1	1	0.99	1	1	1
	Form	1	1	0.99	1	1	1

Table 5.17. – Evaluation of different classification model performance distinguishing between different anthropic percussion marks and carnivore tooth pits.

These initial advances in ML applied to GMM for discerning carnivore agencies has recently been published as Courtenay et al. (2019) in the journal *Palaeogeography, Palaeoclimatology, Palaeoecology*. This is available in Appendix 5.

The training of all these models and evaluation of system performance shows that in total, most algorithms require milliseconds to train (Appendix 2 – Table S6 and Figure S9, Pages 252-253). Predictions on the other hand are even faster (Appendix 2 – Table S7, Page 252), proving how these models are the most efficient classifiers in data science. In general, neither do these models generally use up much

of the CPU either (Appendix 2 – Table S8, Page 254).

The longest training time for any model by far is the case of RF, requiring up to approximately 3 minutes when trained on large amounts of data. As far as system performance is concerned, this is also one of the most costly algorithms using up to about 40% of the CPU at an increased speed of 0.78 GHz. When reducing the number of input variables, this computation can last as little as 44 seconds to train. CPU performance, however, remains the same, proving to be the most costly algorithm to train on data. SVM, another of the most powerful computational models, requires an average of 0.05 seconds to train, proving to be the most powerful and fastest AI model when processing GMM data of any size. In general, SVM also requires the least amount of processing power to train, proving to be one of the most efficient and fastest models for ML. Another powerful model such as C5.0, on the other hand, is almost as fast as SVM to train, lasting at a maximum 0.16 seconds to compute, nevertheless it uses a smaller percentage of the CPU and only pushes the processer 0.61 GHz faster.

In comparison with other models, PLSDA can be seen as one of the most costly models to calculate, lasting up to 15 seconds pushing the CPU the hardest of all the algorithms. KNN and NB on the other hand on the other hand are also fast and easy to programme.

While insignificant differences can be seen in the processing power ($p > 0.05$) and the time taken ($p > 0.05$) to compute models based on the number of outputs, patterns can be seen to show how the number of input variables condition the time taken for a model to be trained. Once the model is trained, an average of 0.01 seconds is required for this model to be used to classify unknown individuals.

Full table of CPU performance is included in Appendix 2 – Table S8, Page 254.

5.8. Deep Learning Classification Models

A total of 4 DNNs were constructed in this study. These consist in 2 DNNs for the binary classification of anthropic and non-anthropic behaviour through circular and linear traces and 2 multiclass classification models to distinguish between different carnivore taxa via the tooth marks presented.

The first binary DNN was used to differentiate between scratches and grazes (Fig. 5.29., Table 5.18.). This model was relatively fast to design, especially considering the relative high separation of groups to start with. In order to construct this model, the final DNN consisted in only 3 layers with 1 input layer (nodes = 10), 1 hidden layer (n = 6), and a final layer with a 'sigmoid' activation. The activation function for all previous layers relied on the popular 'relu' funciton. Considering the high differentiation of both groups in other models, it was considered unnecessary to include dropout layers, cost regulations or any type of weight regularisation. The final model was compiled using the binary crossentropy function, an 'adam' optimizer and using the accuracy metric. The final training was performed using 900 epochs and a microbatch size of 64.

DNN was able to produce 100% classification between samples with a sensitivity of 1, specificity of 1 and a Kappa value of 1 for both validation and testing. The recorded net loss on validation was recorded as 0.016 (Fig. 5.29.).

The second binary DNN was used to differentiate between non-anthropic carnivore agencies and anthropic percussion marks (Fig. 5.30., Table 5.19.). This particular model was fast and easy to design, with only 3 trials before determining the final optimum parameters for this model. The final architecture used to construct this DNN consisted in 5 layers,

reducing the 51 input variables efficiently into 1 output node activated using the 'sigmoid' function. In all previous layers a 'relu' activation layer was used, also employing the use of a kernel constraint weight regularisation algorithm to prevent overfitting via a standardised cost value. The model was compiled using binary crossentropy, an 'adam' optimizer and with metrics set to 'accuracy'. The final number of epochs employed for tuning were 500 with microbatch sizes of 32. This final model produced 100% classification in training, validation and in test sets, with a net loss of 0.004 (Fig. 5.30.). This model presents a significantly more powerful classification than most standard ML algorithms on this type of dataset with a Kappa value of 1 and both sensitivity and specificity values at 1 as well.

While the accuracy can be seen to have peaked early on in the training process (Fig. 5.30.), the training algorithm continued to optimise and reduce loss effectively until the end of the epoch cycle, thus producing no overfitting and proving to be a highly powerful classifier.

The adaptation of this model for multiclass models also proved effective (Fig. 5.31., Table 5.20.). DNN used to study carnivore taxa from tooth pit morphology, regardless of the hyperparameters used, correct classifications in all cases, with a sensitivity and sensibility of 1. Performance on training (acc = 100%, loss = 0.003) and validation data (acc = 100%, loss = 0.002) appears to be highly effective, with perfect loss curves (Fig. 5.31.). Use of this model on test sets also produced up to 100% classification on all accounts.

The construction of this model employed almost the same internal architecture as the DNN trained on this data for binary classification, however, in order to ensure the large number of input variables were reduced effectively, the inclusion of two dropout layers with thresholds of 0.5 and 0.8 respectively, were able to effectively generalise the PC scores for a 4 class identification of different taxa. In this case 'uniform' kernel initialisers were used on each density layer, with a 'relu' activation for all layers minus the final layer where a 'softmax' activation function was employed. A cost function was added to the first layer to reduce possibility of overfitting, while a categorical crossentropy loss functions was used for training. Compiling of the model additionally used an 'adam' optimizer using the 'accuracy' metric. The model was trained using 100 epochs and a microbatch size of 32. Considering the smaller sample sizes for each class, the validation split employed was set to 20% rather than 30%, which is still considered effective.

The fourth DNN was a multiclass model to discern carnivore agency based on tooth scores (Fig. 5.32., Table 5.21). This model was tried and tested a total of 14 times before the optimum hyperparameter configuration was found, reaching a final model accuracy of 99.6%.

This DNN contained the simplest internal architecture with a total of 3 simplistic density layers and no dropout function. Using complex models produced great differences between training and validation sets, with very poor performance on test samples (in some cases 5% classification). The reduction of layers, simplification of their content and the boosting of training epochs was able to increase model performance by 42.86% in the final model. For this DNN, the first two layers simply consisted of a 'relu' activation function and the final layer used a 'softmax' activation function to define the 4 class groups. The model was compiled using categorical crossentropy for loss, the 'adam' optimiser and with metrics set to 'accuracy'. The final training of the DNN was where the majority of hyperparameter tuning came into play. Low epoch values did not provide the model enough time to reach an optimal loss value, while the batch size greatly affected the time needed for the model to learn the data. The final model was trained using 2000 epochs and a microbatch size of 32, reaching very effective classifications on all accounts (train acc/loss = 99.86/0.01, val = 99.52/0.03, test acc = 99.6) (Fig. 5.32.).

The final binary DNN classifier constructed was used to study anthropogenic cut marks and trampling marks (Fig. 5.33., Table 5.22.). This DNN required a long series of hyperparameter optimisation trials (n = 51) before a final model could be determined (a summary of trials have been included in Appendix 2 – Table S9, Pages 255 to 256). Initial trials began with a model accuracy of 70.3% and, in most cases, overfitting was still a considerable factor even at these low accuracy rates. Nevertheless, a final model was obtained with an architecture that consisted in 6 layers; 5 standard neural layers and 1 dropout layer. While the number of inputs in this case are considerably lower, case studies have seen how relying on simple architectures are likely to produce a bottleneck effect, compressing and reducing the variables too quickly and producing greater chances of under or overfitting. Because of this, one of the trials consisted in providing the model with an intermediate layer of a greater density to prevent this phenomenon from occurring. Once observing that this process produced a boost in accuracy, this layout was kept for all following layers. In order to prevent this additional layer from producing over generalisation of the data, a dropout layer with a constraint of 0.5 was included in layer 4. The position and number of different dropout layers were tried and tested with the most effective position and threshold in-between layers 3 and 5.

Activation parameters for the all initial layers used the 'relu' function, with a 'sigmoid' final activation function for the binary output. After observing overfitting on multiple accounts, an additional kernel constraint weight regularisation algorithm was used to correct this issue. The model was compiled using binary crossentropy, an 'adam' optimiser and 'accuracy' set for metrics. The final training process used 900 epochs and a microbatch size of 64.

The highest classification model wasn't achieved until the kernel initialiser was removed from this model. This small variation finally achieved the 100% classification range on all training, validation and test samples (loss train = 0.019, validation = 0.001). Model performance can be considered highly powerful with Kappa, sensitivity and specificity value of 1.

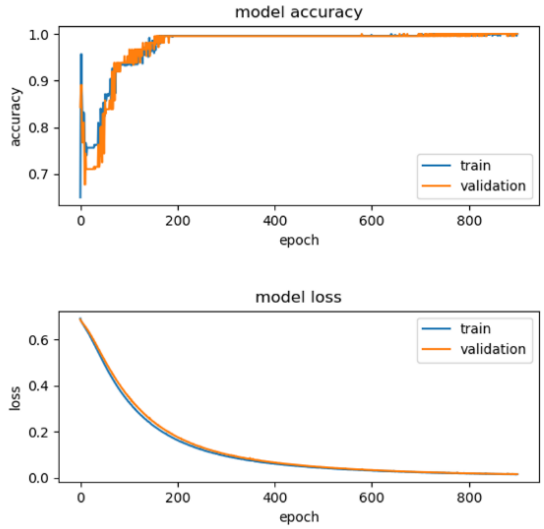


Figure 5.29. – Accuracy and loss scores plotted against epochs during the training process of a DNN that differentiates between scratches and grazes.

	Layer (Type)	Output Shape	Number of Parameters
1	Dense	10	110
2	Dense	6	66
3	Dense	1	7

Table 5.18. – Final DNN internal architecture used to create a classification model that distinguishes between scratches and grazes.

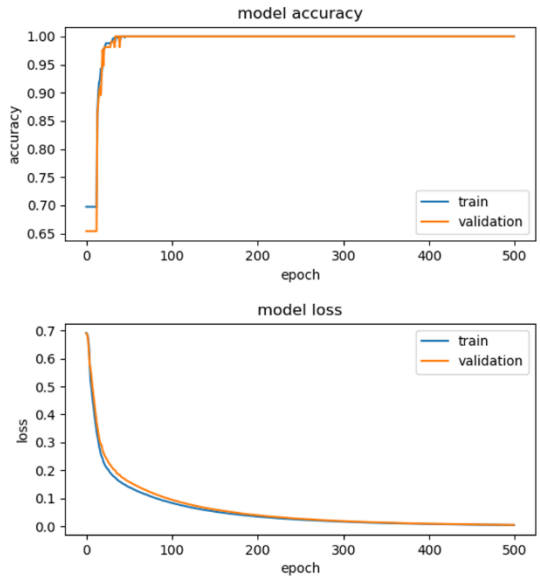


Figure 5.30. – Accuracy and loss scores plotted against epochs during the training process of a DNN that differentiates between percussion marks and tooth pits.

	Layer (type)	Output Shape	Number of Parameters
1	Dense	51	2652
2	Dense	25	1300
3	Dense	10	260
4	Dense	5	55
5	Dense	1	6

Table 5.19. – Final DNN internal architecture used to create a classification model that distinguishes between percussion marks and tooth pits.

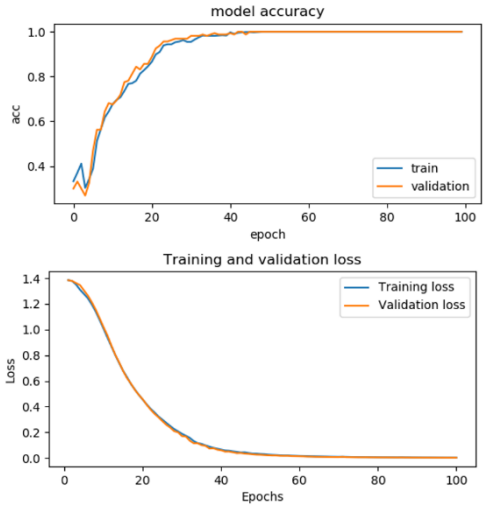


Figure 5.31. – Accuracy and loss scores plotted against epochs during the training process of a DNN that differentiates between different taxa based on tooth pit morphology.

	Layer (type)	Output Shape	Number of Parameters
1	Dense	51	2652
2	Dropout	51	0
3	Dense	25	1300
4	Dropout	25	0
5	Dense	15	390
6	Dense	4	64

Table 5.20. – Final DNN internal architecture used to create a classification model that distinguishes between different taxa based on tooth pit morphology.

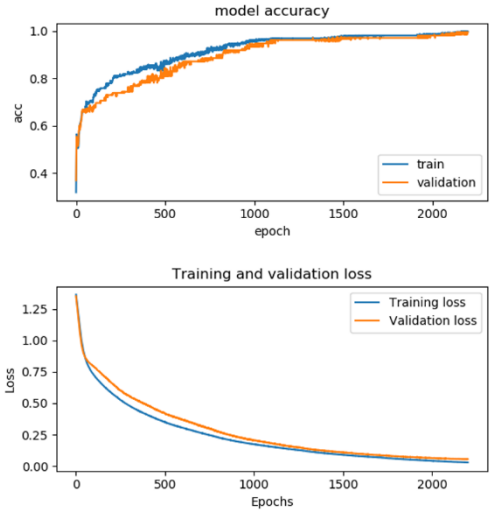


Figure 5.32. – Accuracy and loss scores plotted against epochs during the training process of a DNN that differentiates between different taxa based on tooth score morphology.

	Layer (type)	Output Shape	Number of Parameters
1	Dense	10	110
2	Dense	8	88
3	Dense	4	36

Table 5.21. – Final DNN internal architecture used to create a classification model that distinguishes between different taxa based on tooth score morphology

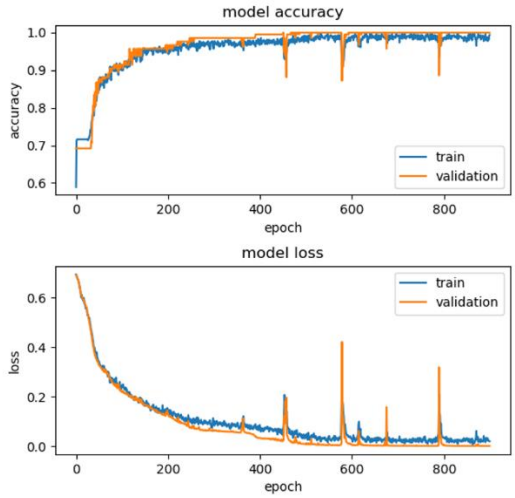


Figure 5.33. – Accuracy and loss scores plotted against epochs during the training process of a DNN that differentiates between cut marks and trampling marks.

	Layer (type)	Output Shape	Number of Parameters
1	Dense	10	110
2	Dense	20	220
3	Dense	10	210
4	Dropout	10	0
5	Dense	6	66
6	Dense	1	7

Table 5.22. – Final DNN internal architecture used to create a classification model that distinguishes between cut marks and trampling marks.

Training time for DL models is entirely dependent on the epochs and batch size used during training, therefore comparisons of the times taken to construct these models is difficult. Nevertheless, the training time for each model has been summarised in Appendix 2 – Table S10, Page 257. As can be seen, the time taken to train these models is still much faster than RF, and neither is this a very costly method to train as far as time is concerned. The more difficult part of this type of classification exists in the programming of the model and tuning of hyper parameters.

As for processor performance, on average, DL uses up to 83 +/- 10% of the CPU and pushes the speed a further 0.49 GHz. This can thus be considered the model that requires the most amount of processing power to train, however considering the speed, accuracy and complexity of the model, it can still be considered one of the most efficient classifiers.

For the purpose of providing a visual example of how a DNN is composed internally, a visual example of the trampling vs cut mark model has been provided in Fig. 5.34.

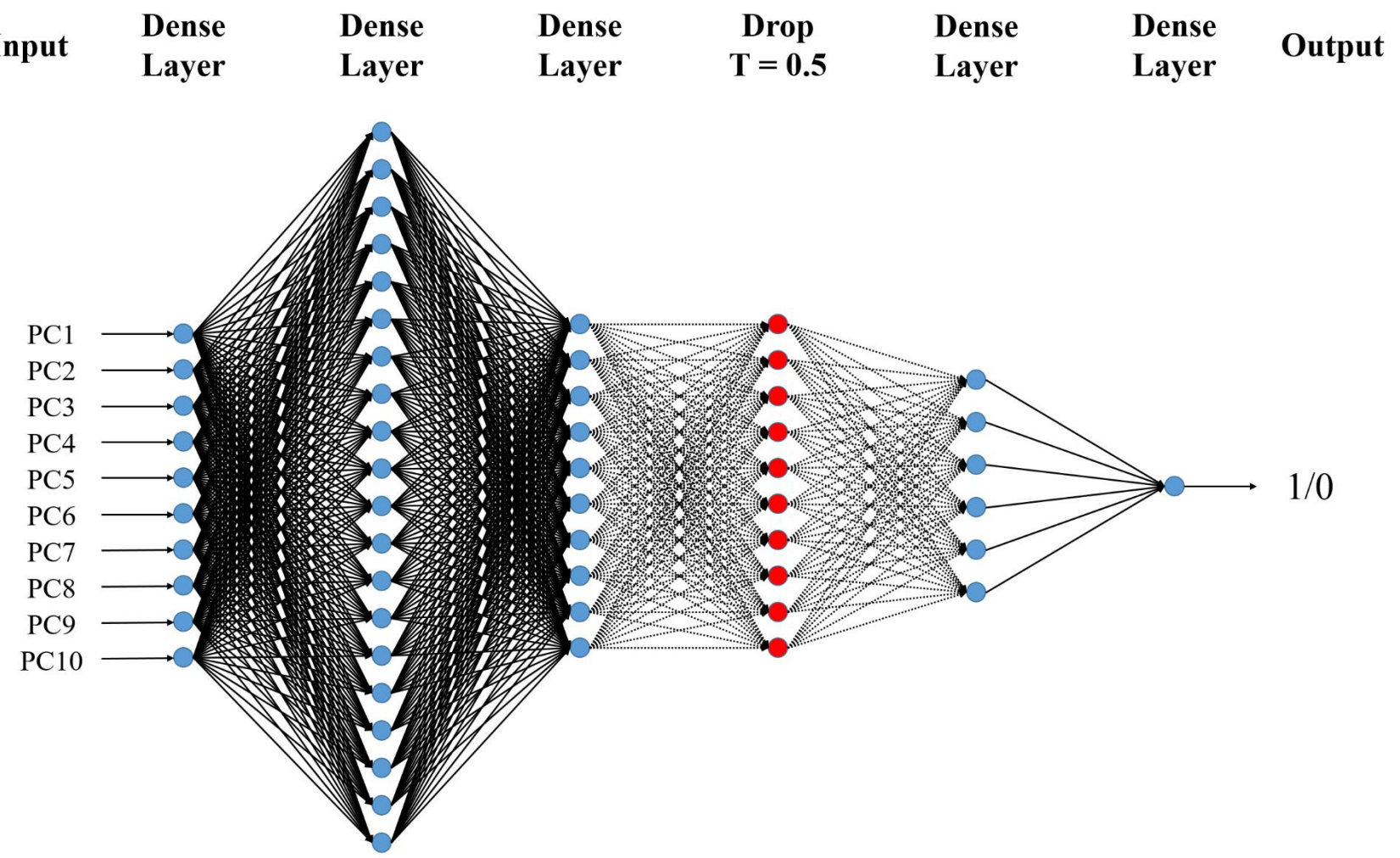


Figure 5.34. – Example of Deep Learning Neural Network used to process Geometric Morphometric data of linear traces, producing a final binary classification (1 = Anthropic/0 = Natural)

5.9. The Final Models

Of all ML and DL algorithms, DNN, SVM, RF and C5.0 were instantly recognised as the best performing models when trained on GMM data. All 4 of these models were finally extrapolated and prepared to be used for the classification of unknown archaeological data from the FLK-West archaeological assemblage.

Of these 4 models, all present a total loss of approximately 0.01, thus describing the general confidence of each models classification as 98.97% (Table 5.23.). The most confident model is the RF classification of scratch and graze marks, with an average of 99.97% confidence when assigning class probability. The least confident model is found in the DNN's classification of carnivore taxa through tooth scores, with an average of 93.3% confidence of each classification made. Considering how this is the worst performing model (accuracy = 99%, kappa = 0.99) of all the "top performing models", this result is expected. Excluding DNN for scores, the highest loss value is found in DNN's classification of scratches and grazes, with an average confidence of 97.3% for each class assigned.

Nevertheless, all models are very powerful classifiers and present a significant advance from the average LDA test used in classical GMM approaches.

	Net Loss				
	Scr Vs Gr	**CM vs Tr**	**PM vs TM**	**Pit Taxa**	**Score Taxa**
SVM	0.0057	0.0059	0.0039	0.0160	0.0190
RF	0.0003	0.0068	0.0150	0.0019	0.0150
C5.0	0.0014	0.0014	0.0014	0.0043	0.0043
DNN	0.0270	0.0039	0.0037	0.0025	0.0670

Table 5.23. – Table presenting the loss values obtained when testing model performance on test sets for each of the ML and DL models when prepared using GMM data. In order of appearance: Scratch vs Graze, Cut Mark vs Trampling Mark, Percussion Marks vs Tooth Marks, Taxa from carnivore Tooth Pits and Taxa from carnivore Tooth Scores.

5.10. FLK-West

A total of 25 linear marks and 4 circular marks were found and located on the FLK-West faunal assemblage. 24 of these linear marks could be processed using Courtenay et al. (2017)'s landmark model and all 4 circular marks could be processed using Aramendi et al. (2017a)'s landmark model. The excluded linear mark was removed from the sample considering how not all 13 landmarks could clearly be defined across the entire groove.

Through general inspection of each bone with the aid of ESEM, other taphonomic processes identified over these remains included heavy water abrasion and polishing on almost all bones (Fig. 5.35.). In most cases, this rounding produced by fluvial abrasion can appear on both the edges of the bone's fracture planes as well as the edges of the marks present. Additional biochemical alterations were also identified on some bones (Fig. 5.36.), as well as one linear mark that was confirmed upon further inspection to be a modern mark, most likely product of excavation (Fig. 5.37.).

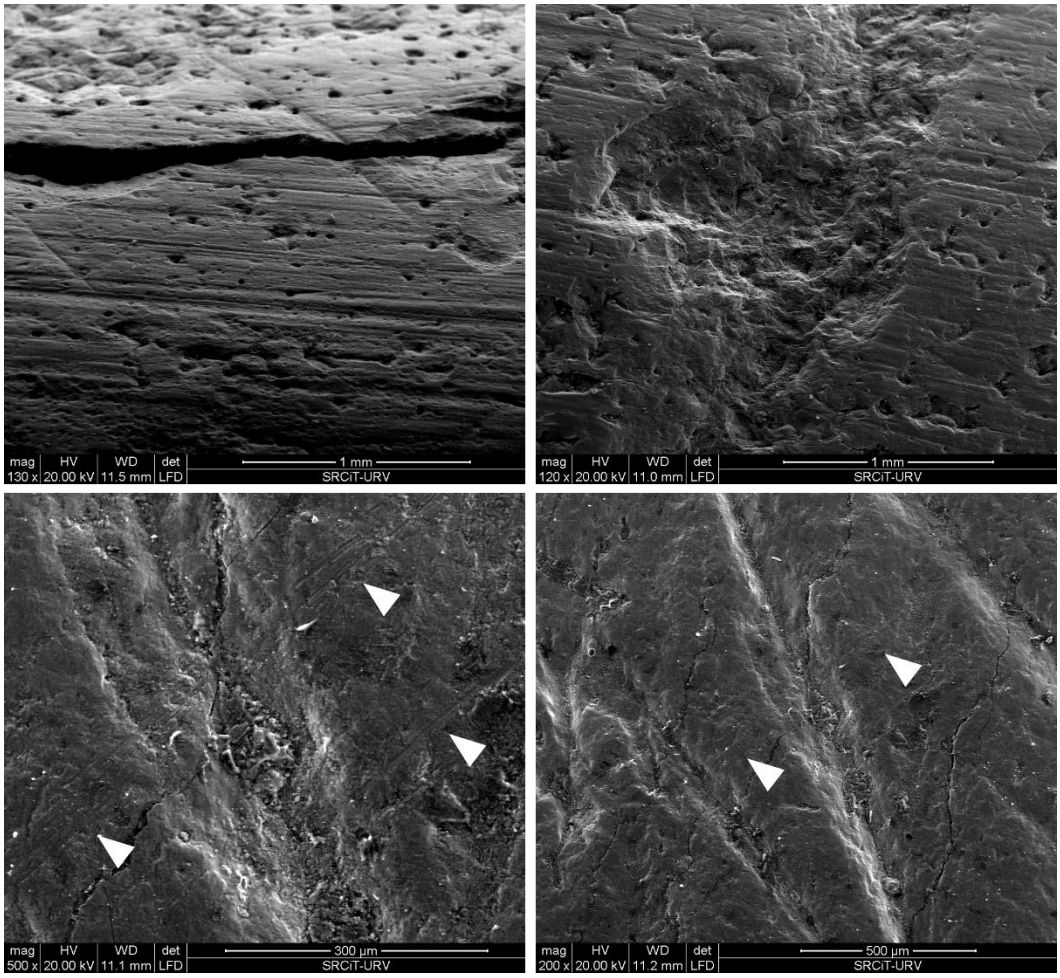


Figure 5.35. – ESEM photos of multiple bones presenting heavy and moderate alteration of bone surfaces by fluvial abrasion. The top two images present very heavy abrasion, with the top right mark being completely distorted by this taphonomic process. The bottom two photos present a lower degree of fluvial alteration, nevertheless additional striae can be observed overlapping the original BSMs (marked with white arrows) and the edges of these grooves appear to have been rounded by said process.

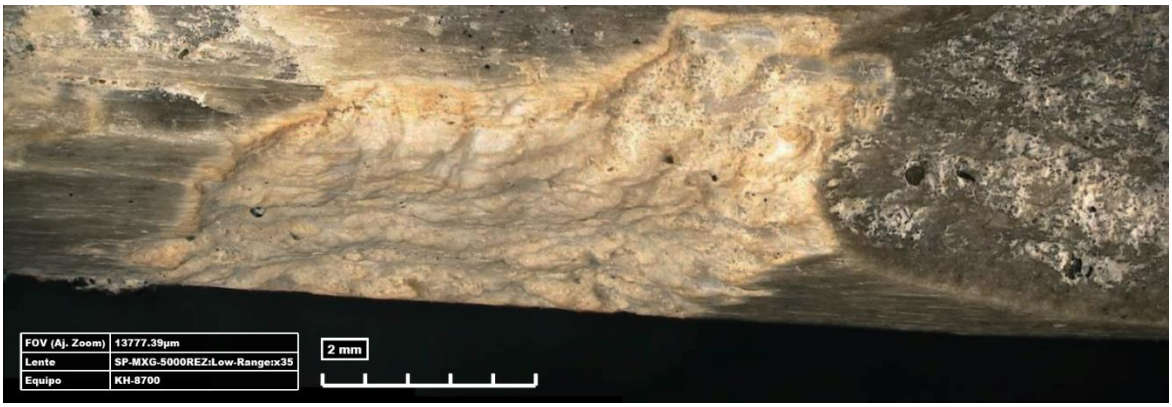


Figure 5.36. – Biochemical alterations observed using the HIROX

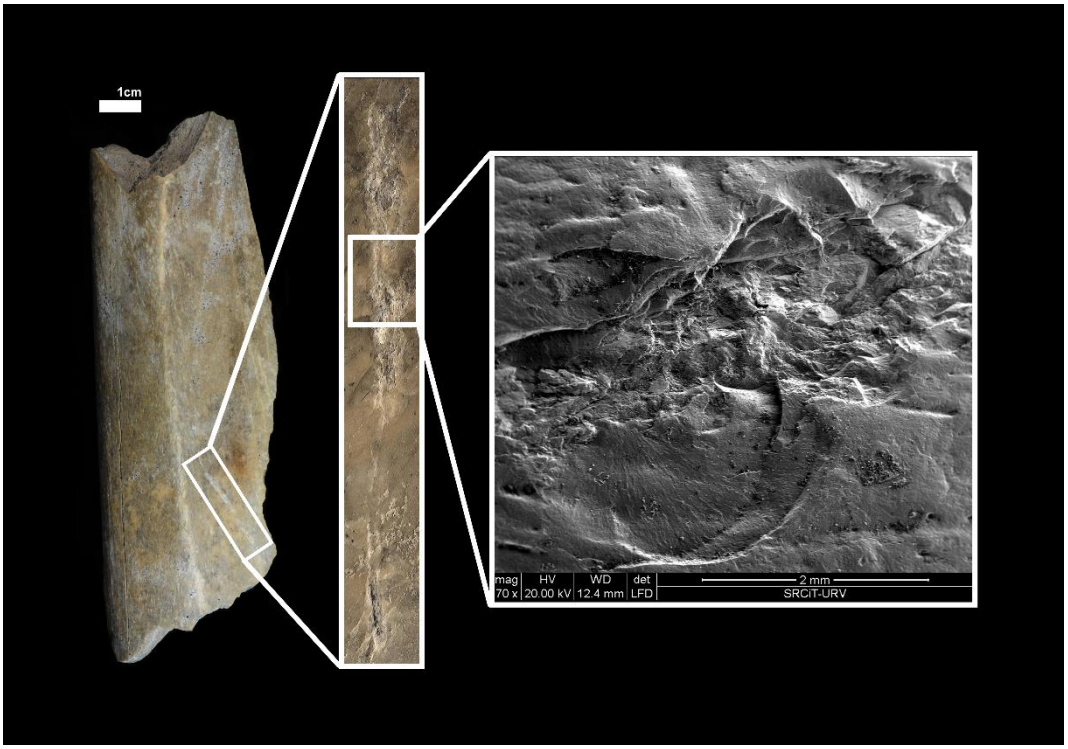


Figure 5.37. – Additional damage to bone produced while bone is dry, most probably produced during the excavation process.

While most marks present some of the diagnostic features described by Domínguez-Rodrigo et al. (2009a), the effect of water abrasion has damaged most of these traces on multiple accounts. Nevertheless, neither are these diagenetic and biostratinomic processes strong enough to have eliminated all features of these traces. Some marks present significant degrees of abrasion, yet internal microstriations are still observable in some cases at sporadic points along the incision. Because of this, these marks are incomparable with case studies such as that of Barranc de la Boella (Pineda et al., 2014), where biostratinomic processes appear to have eliminated internal microstriae in their entirety. This single feature is strong enough to condition the entirety of these statistically analysed qualitative results. Traditional comparison of cut mark features for this site are, therefore, futile and not likely to provide a clear comparison with any of the experimental samples from a qualitative perspective (Appendix 2 – Figure S10, Page 257).

Needless to say, the small sample size is also a considerate conditioning factor. Considering these observations, however, it is clear the FLK-West suffered a significant degree of biostratinomic or diagenetic alterations that affected the preservation rate of bone cortical surfaces, however these alterations are not considerable enough to have completely eliminated some of their diagnostic features.

GMM analysis of BSMs and testing of AI algorithms on archaeological data proved successful in providing robust classifications for all marks, however, not all results can be considered conclusive, revealing interesting information that can be used to orientate further investigation into BSMs and this site in particular.

Reflecting on the performance of each algorithm (Appendix 2 – Tables S11 to S14, Page 258), C5.0 overclassified all traces with very low confidence intervals for each classification. This proves that this particular algorithm most likely overfitted on both testing and training data, and was unable to generate accurate results when applied to archaeological data. In the case of RF a similar phenomenon

can be observed. RF proved average results, with a mixed classification of many marks and relatively low classification confidence boundaries. At best, the model was 88% sure of one of the classifications made on a linear trace, yet on the other hand, the worst identification of a BSM was seen in the case of one of the circular depressions, with 50.4% confidence in classing this mark as a tooth mark.

SVM and DNN, however, were very powerful models on all accounts, reaching up to 100% confidence with some classifications, and an average of >90% confidence with almost all others. Needless to say, in some cases neither model could come to an agreement as to what classed label to provide the archaeological data, thus proving that the effect of water on the morphology of these traces to be significant and need of further investigation. Interestingly, the marks that proved the most complicated to classify where all found on the same bones, with the highest degree of insecurity appearing on bones with greater presence of fluvial abrasion. This arguably shows how the taphonomic context for these traces is a significant factor in their morphological characterisation.

Nevertheless, in cases where equifinality was still present and the models were not able to converge, the most confident prediction value was taken (Table 5.24. & 5.25.). For example, if SVM classed the mark as a cut mark with 94.29% confidence, and DNN disagreed with 96.74%, the final decision would be to label this mark as a trampling mark.

Final Class	Classifier	Confidence (%)	Trampling Class	Classifier	Confidence (%)
Trampling	SVM	96.74	Graze	Both	99.99
Trampling	Both	100.00	Graze	Both	100.00
Trampling	Both	99.25	Graze	Both	100.00
Trampling	Both	99.99	Graze	Both	100.00
Cut Mark	Both	98.76			
Cut Mark	Both	99.95			
Cut Mark	Both	99.98			
Trampling	Both	82.00	Scratch	Both	91.47
Trampling	Both	99.93	Graze	Both	100.00
Trampling	Both	99.95	Graze	Both	100.00
Cut Mark	Both	99.93			
Cut Mark	Both	99.94			
Cut Mark	Both	99.97			
Cut Mark	Both	97.31			
Trampling	Both	99.93	Graze	Both	100.00
Cut Mark	Both	99.93			
Cut Mark	Both	99.99			
Cut Mark	SVM	99.47			
Cut Mark	SVM	99.93			
Cut Mark	DNN	100.00			
Cut Mark	Both	99.85			
Trampling	Both	99.99	Graze	Both	100.00
Cut Mark	DNN	99.27			
Cut Mark	SVM	99.19			

Table 5.24. – Final classifications for FLK-West linear traces using SVM and DNN classifiers

Final Class	Classifier	Confidence (%)	Taxa	Classifier	Confidence (%)
Percussion	DNN	98.79			
Tooth	DNN	100.00	Lion	DNN	99.99
Tooth	DNN	100.00	Lion	DNN	85.11
Tooth	DNN	100.00	Lion	DNN	99.67

Table 5.25. – Final classifications for FLK-West circular depressions using SVM and DNN classifiers

GMM analysis on linear traces produced a total of 32 PC scores in PCA, with the first 2 PC scores representing a combined cumulative variance of 55.14% of the sample. Upon analysing shape variation across PC scores using thin plate splines, PC1 can be primarily characterised by the width of the mark while PC2 represents the depth.

Of the 25 linear marks observed, 15 of these were classed at cut marks, while 9 as trampling marks (Fig. 5.38. and 5.39.). In general, the FLK-West cut marks lean closer to the shallower marks on PC2, possibly indicating the more shallow nature produced by fluvial abrasion, while the trampling marks are much more varied. MANOVA was unable to identify significant associations of classed archaeological incisions to either of the experimental samples (Cut Mark $p = 0.001$ and Trampling $p = 0.003$).

Carrying out a CVA from newly labelled data produced inconclusive results (Fig. 5.40.). In both cases, the archaeological data was positioned away from both experimental samples, indicating they may be something quite different altogether. Additionally, Mahalanobis and Procrustes distance calculations provided contradictory numerical values (Table 5.26.). Needless to say, the archaeological samples are too small to be considered significant in this type of analysis, and until a larger sample is obtained, this data should be discarded from this study.

Once classed, the trampling marks were further subjected to classification to discern their trampling type. This particular GMM analysis produced a similar number of PC scores to the other linear model (Fig. 5.41.), yet differentiation between samples is obviously much clearer. Of the 9 trampling marks identified, 8 were classified as grazes with over 99.9% confidence on all accounts. The only mark classed as a scratch seemed a little less confident, nevertheless MANOVA results displayed significant similarities thus supporting this classification ($p = 0.336$). CVA results could not be assessed considering issues in the archaeological sample size.

Analysis of circular depressions presented a complex PCA with up to 51 PC scores presenting high degrees of overlapping on all accounts. Nevertheless, considering the efficiency of DNN on sorting through this kind of data, DNN was able to produce an average of 99.7% confidence when classifying three of these pits as tooth marks and 1 pit as a percussion mark (Fig. 5.42.). SVM on this account proved to be poor classifier when differentiating between groups. Subject to further analysis (Fig. 5.43. and Fig. 5.44.), these three tooth marks were classed as lion tooth marks with an average confidence of 94.9% when providing this classification.

On all accounts these classes were considered significant when referring to MANOVA results, considering both classification of the percussion mark ($p = 0.218$), and the association of all three tooth marks with the activity of lions ($p = 0.257$). Nevertheless, due to small sample size, CVA could not be performed.

DNN and SVM have thus proved highly powerful classifiers, presenting clear advantages for processing high dimensional overlapping data such as that produced by taphonomic GMM. Nevertheless, the clear presence of fluvial abrasion is a topic that needs to be addressed when using this approach to confront the rest of the taphonomic register present at FLK-West.

The following pages present all the final documentation of taphonomic traces (Fig. 5.38. to 5.44.) with visual documentation for the traces and their associated classification (Fig. 5.45. to 5.52.).

			Archaeological			
			Cut		Trampling	
			D	*p*	*D*	*p*
Experimental	Cut	Mahalanobis	4.4187	< 0.0001	5.2939	< 0.0001
		Procrustes	0.385	< 0.0001	0.0541	< 0.0001
	Trampling	Mahalanobis	2.9745	< 0.0001	5.412	< 0.0001
		Procrustes	0.0113	0.821	0.0611	< 0.0001

Table 5.26. – Mahalanobis and Procrustes distances with associated P Values of the classed archaeological marks with the experimental marks after CVA.

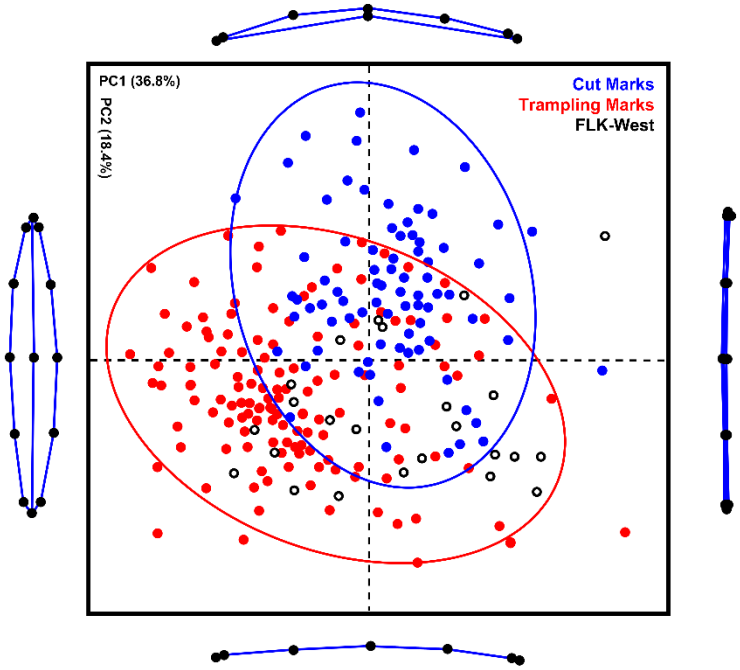


Figure 5.38. – PCA in shape space of unclassed linear marks and 95% confidence intervals for experimental samples in comparison with archaeological marks from FLK-West. Variances in shape are represented on either extremity of their corresponding PC score axis.

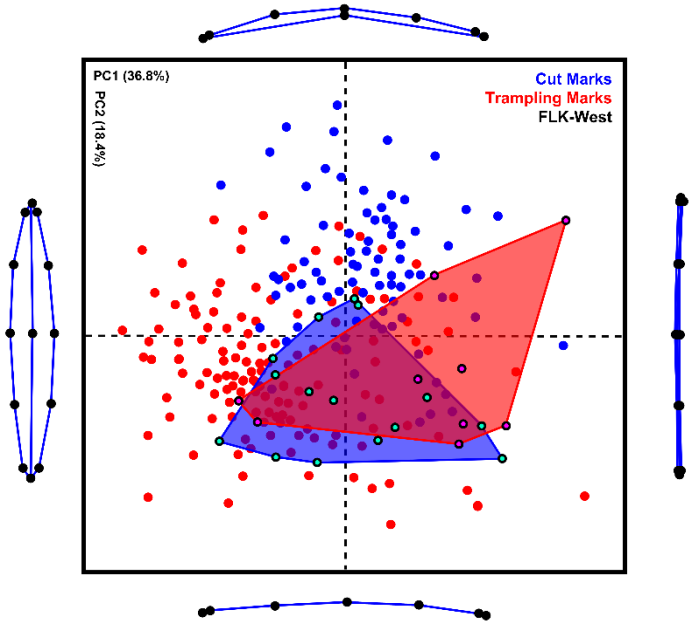


Figure 5.39. – PCA in shape space of classed linear marks using ML and DL algorithms from FLK-West with convex hulls highlighting archaeological marks and their classes. Variances in shape are represented on either extremity of their corresponding PC score axis. Classed mark colours: Light blue = FLK-West marks classed as cut marks. Pink = FLK-West marks classed as trampling marks

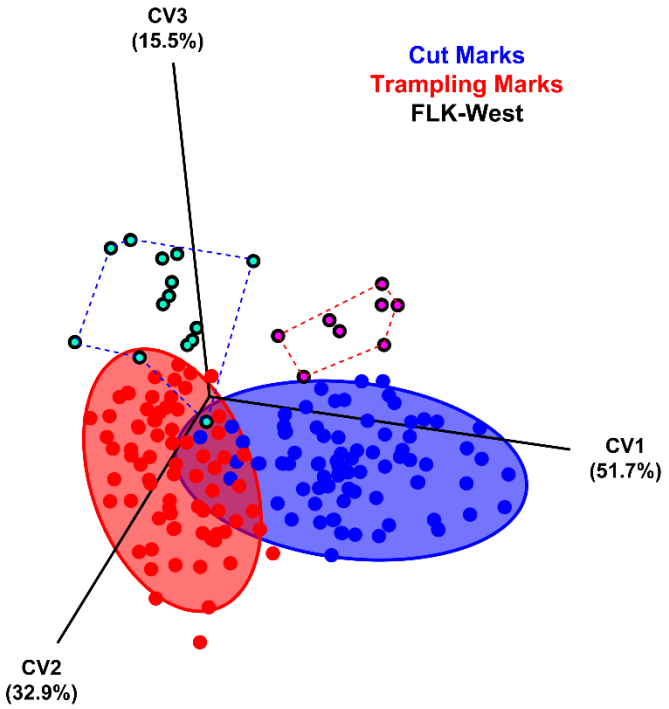


Figure 5.40. – 3D scatter plot of CVA comparing classed linear marks from FLK-West with corresponding experimental samples. 95% confidence intervals have been calculated and included for the experimental samples alongside convex hulls delimiting the different classed archaeological marks. Classed mark colours: Light blue = FLK-West marks classed as cut marks. Pink = FLK-West marks classed as trampling marks

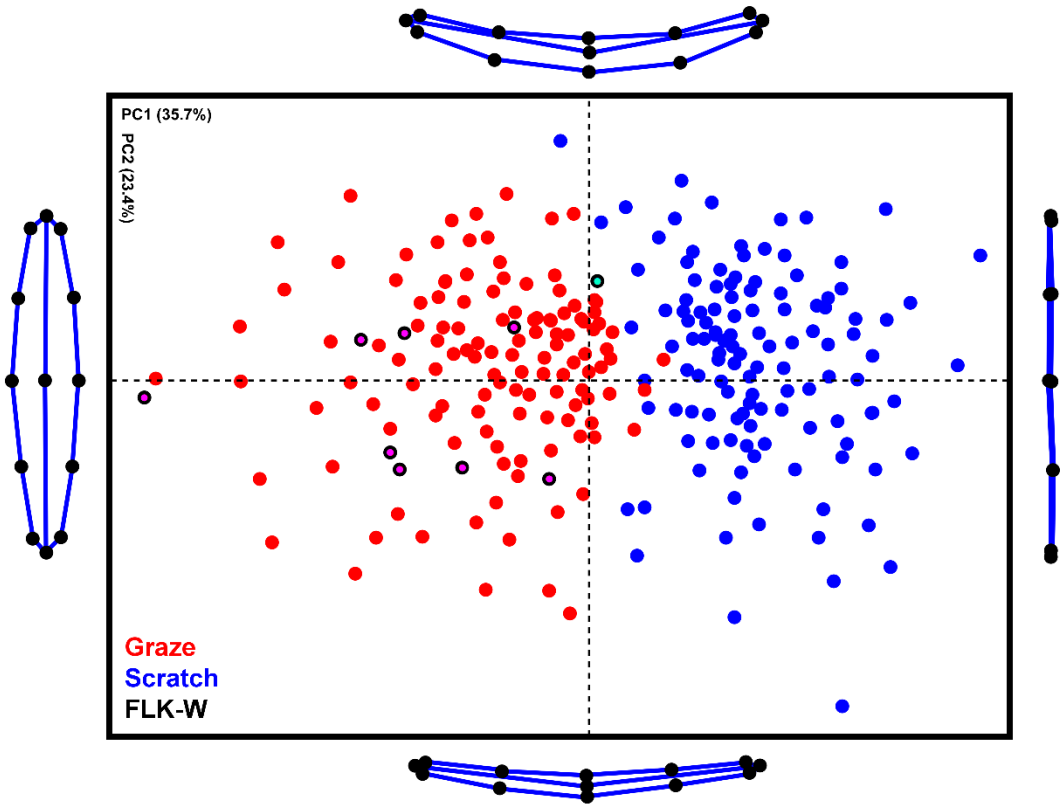


Figure 5.41. – PCA in shape space of classified types of FLK-West trampling marks using ML and DL algorithms. Variances in shape are represented on either extremity of their corresponding PC score axis. Classed mark colours: Light blue = FLK-West marks classed as scratch marks. Pink = FLK-West marks classed as graze marks

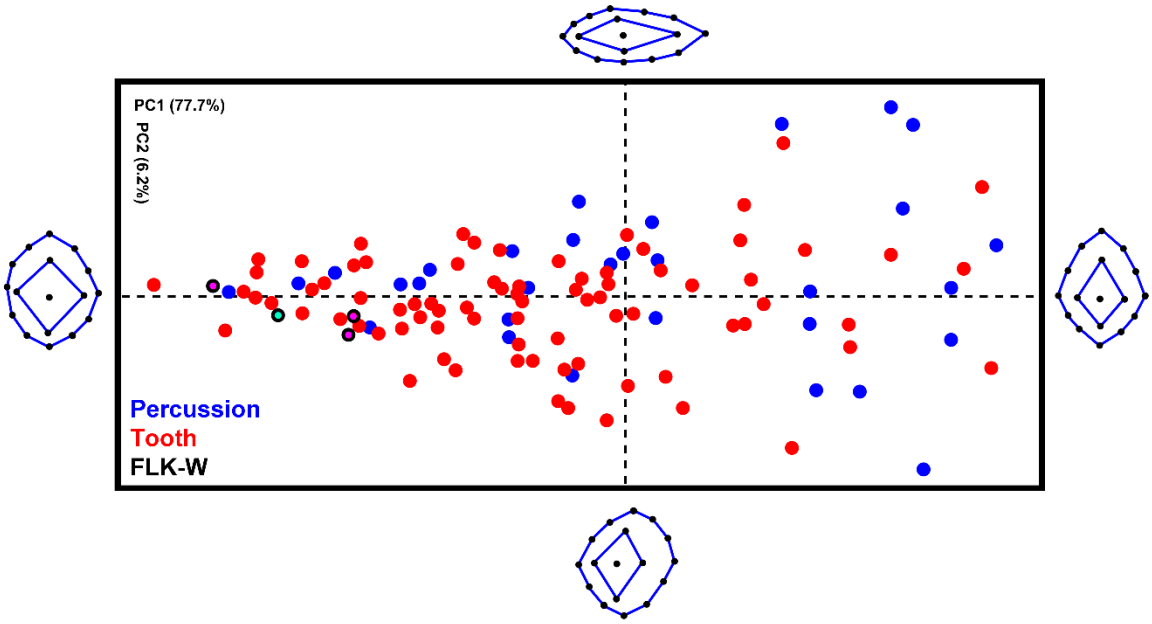


Figure 5.42. – PCA in form space of classified FLK-West circular depressions using ML and DL algorithms. Variances in form are represented on either extremity of their corresponding PC score axis. Classed Mark Colours: Light blue = FLK-West marks classed as percussion marks. Pink = FLK-West marks classed as tooth marks

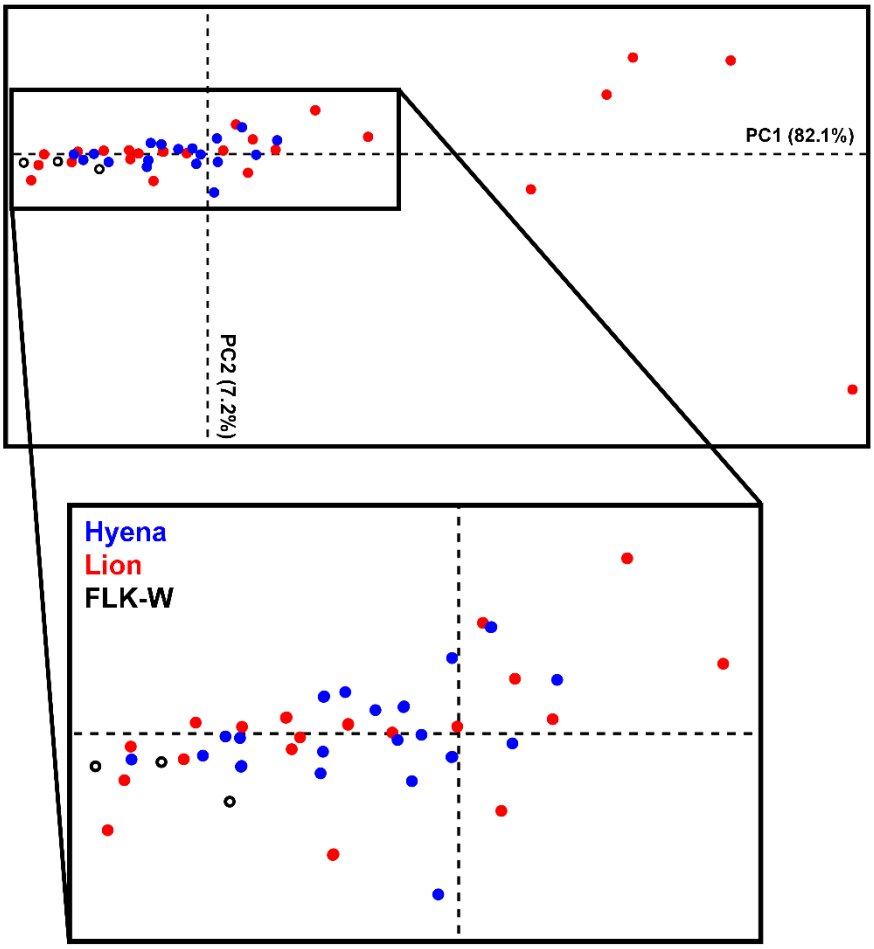


Figure 5.43. – PCA in form space of unclassified FLK-West tooth marks.

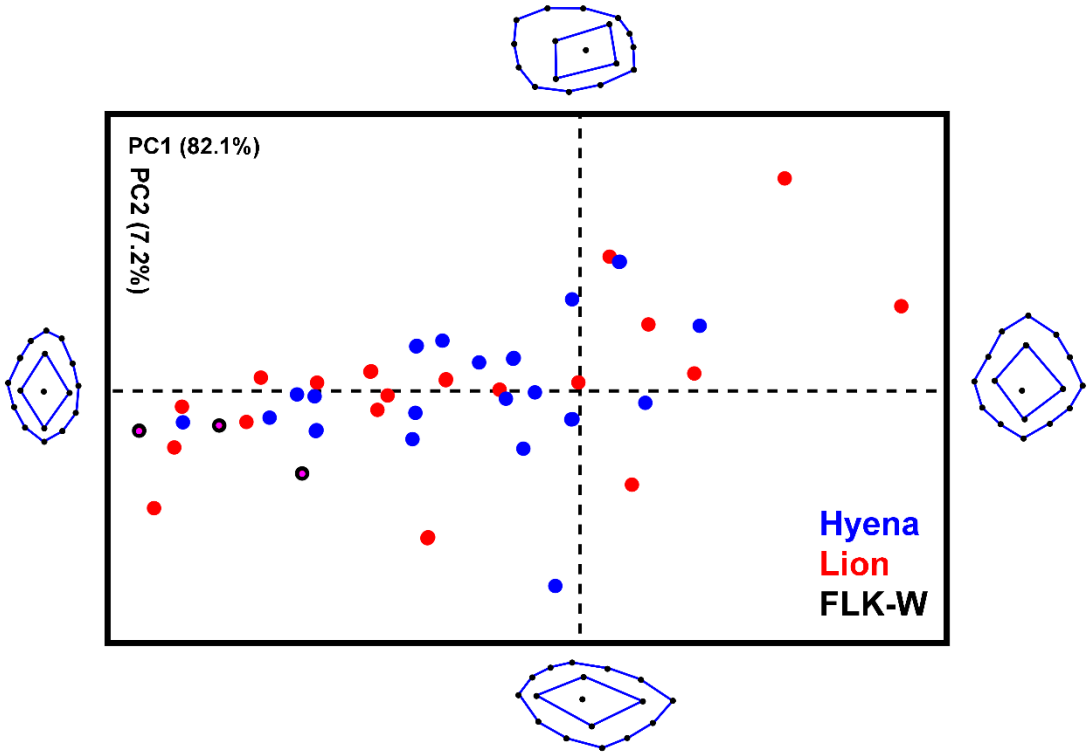


Figure 5.44. – PCA in form space of classified FLK-West tooth marks using ML and DL algorithms. Variances in form are represented on either extremity of their corresponding PC score axis. Classed Mark Colours: Light blue = FLK-West marks classed as hyena. Pink = FLK-West marks classed as lion

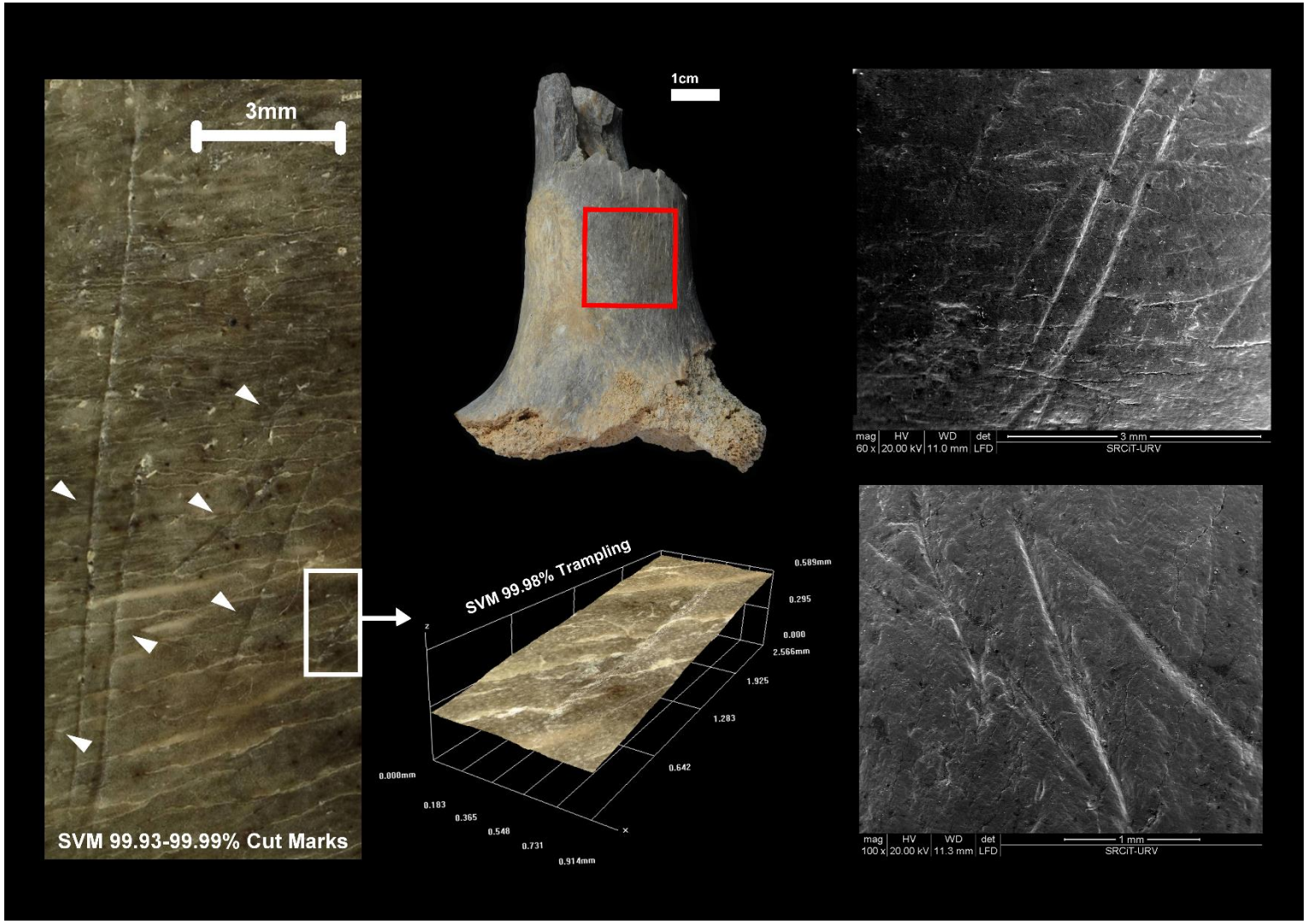


Figure 5.45. – Documentation using ESEM and 3D Digital Microscopy of taphonomic traces found on the distal epiphysis of a bovid femur from the FLK-West archaeological assemblage alongside their classified labels. White arrows in far left photograph indicate presence of cut marks.

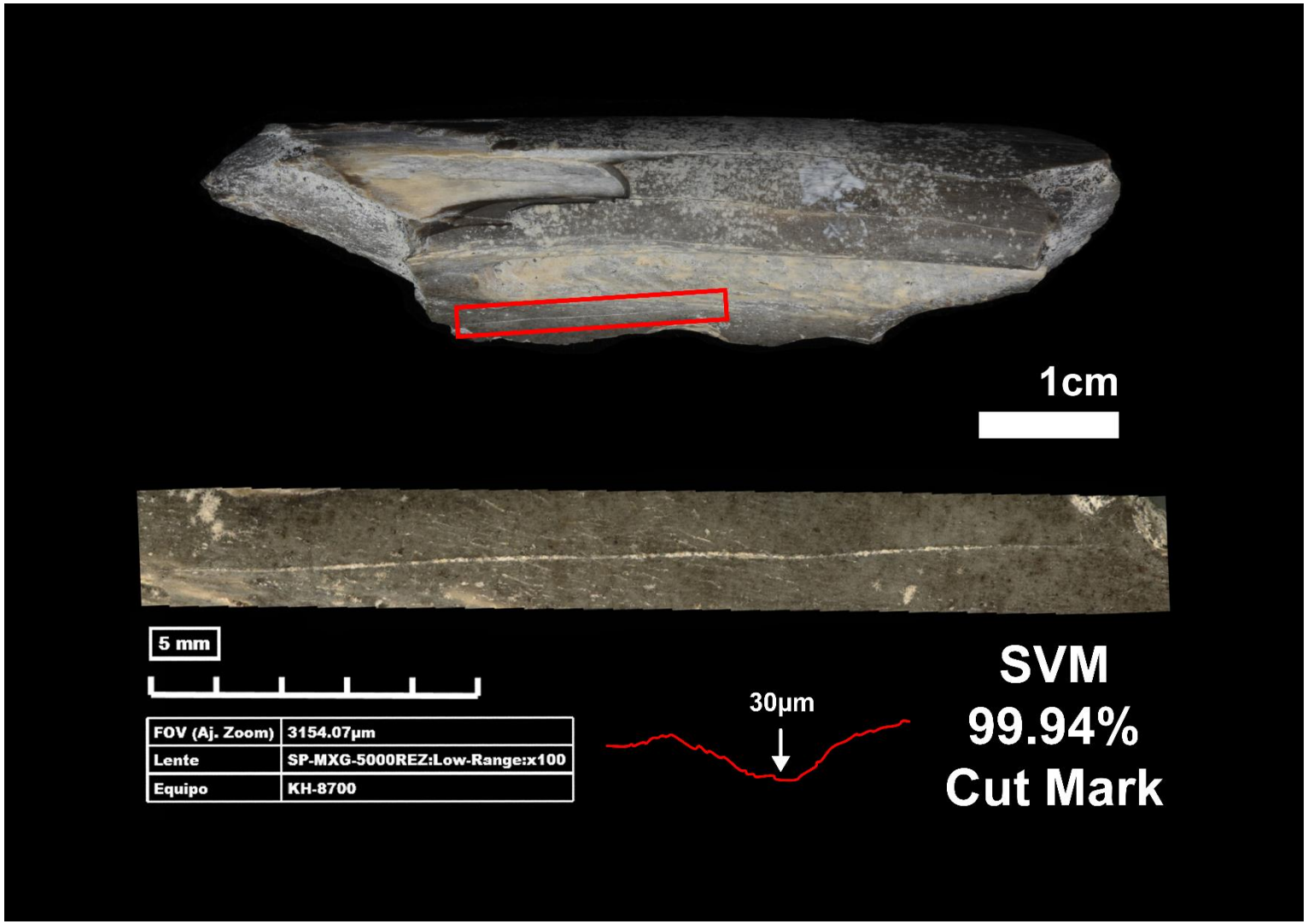


Figure 5.46. – Documentation using 3D Digital Microscopy of taphonomic traces found on the diaphysis of a bovid radius from the FLK-West archaeological assemblage alongside their classified labels.

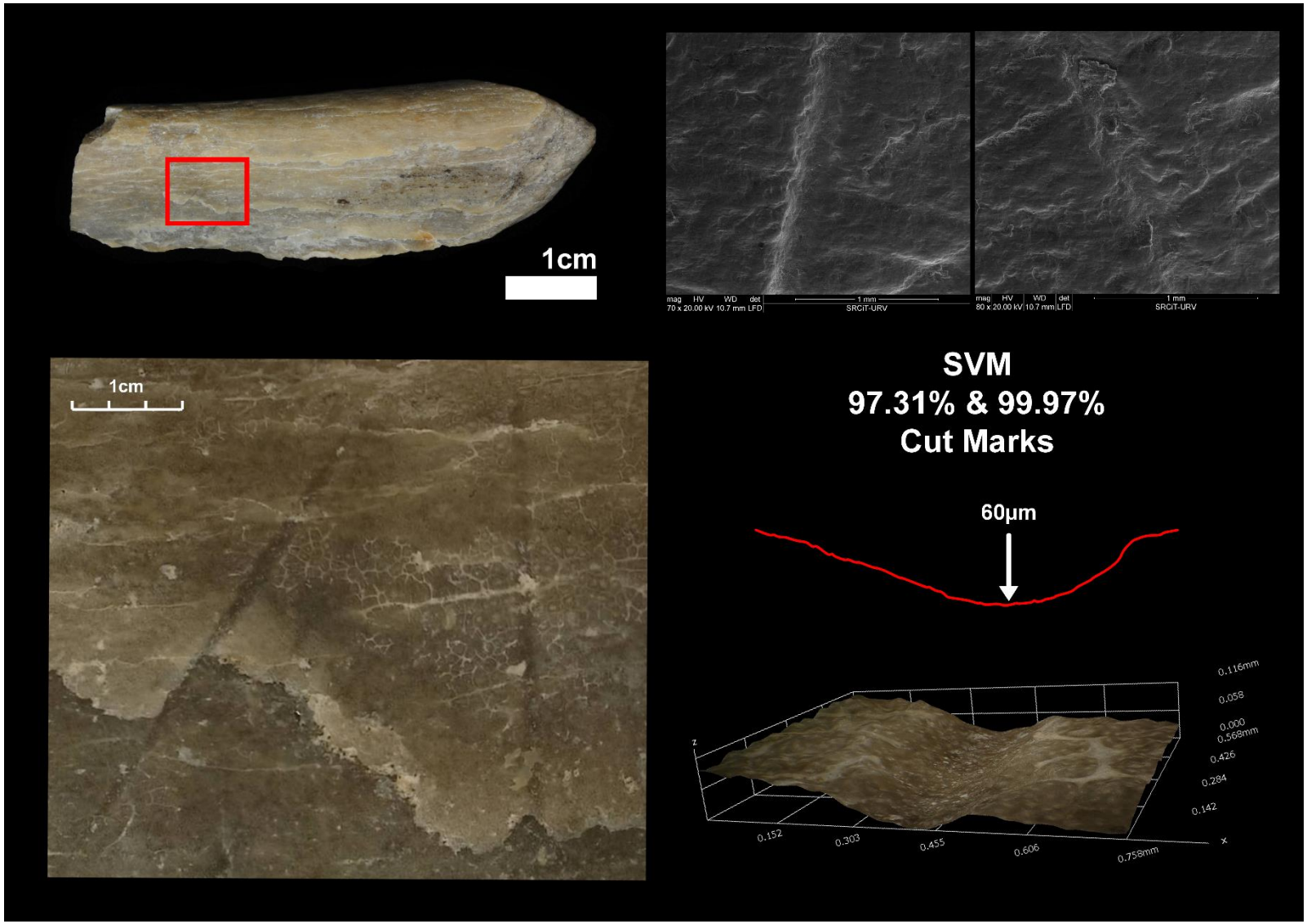


Figure 5.47. – Documentation using ESEM and 3D Digital Microscopy of taphonomic traces found on the diaphysis of a radius (from an indeterminable species) from the FLK-West archaeological assemblage alongside their classified labels.

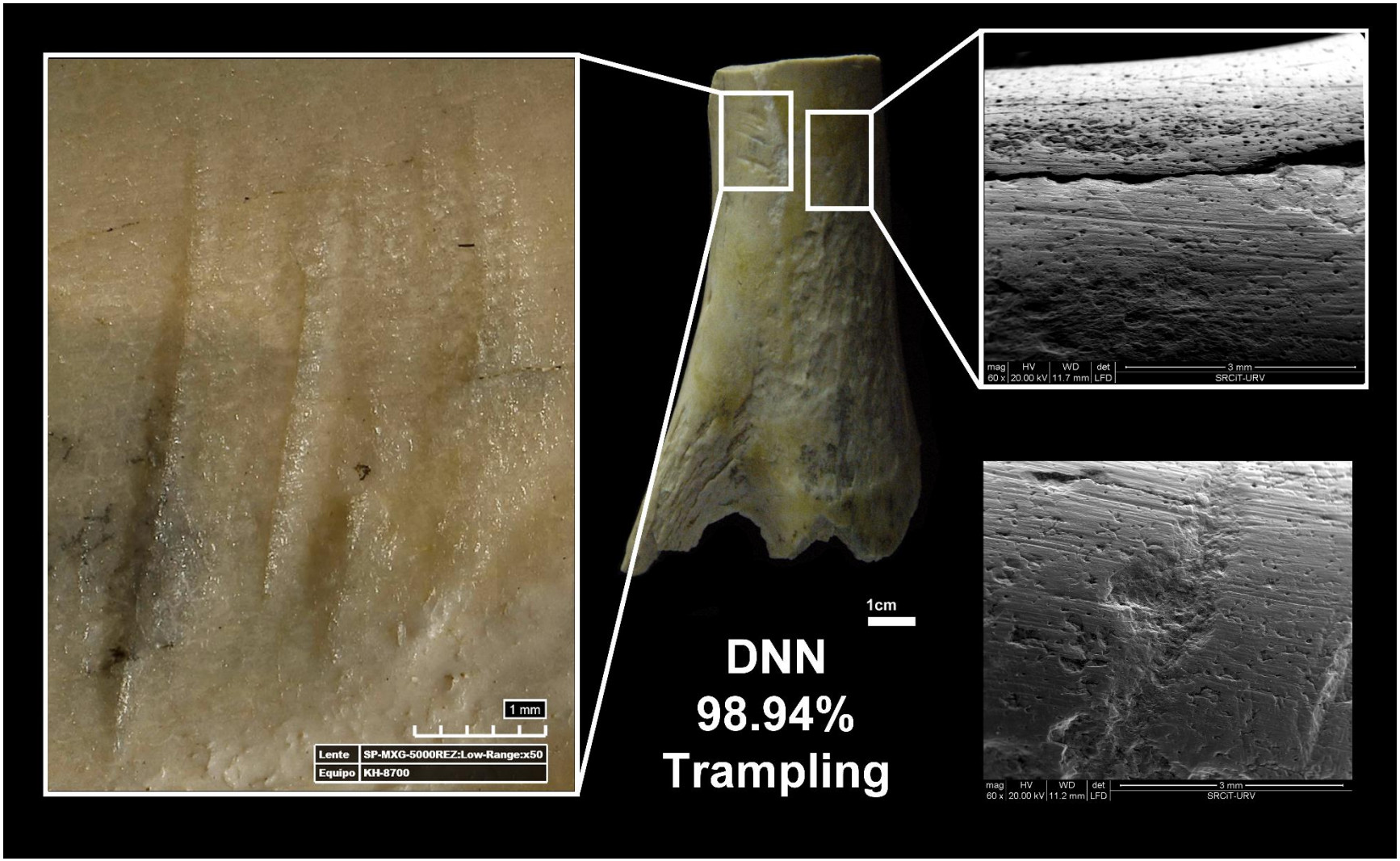


Figure 5.48. – Documentation using ESEM and 3D Digital Microscopy of taphonomic traces found on the distal epiphysis of a bovid humerus from the FLK-West archaeological assemblage alongside their classified labels.

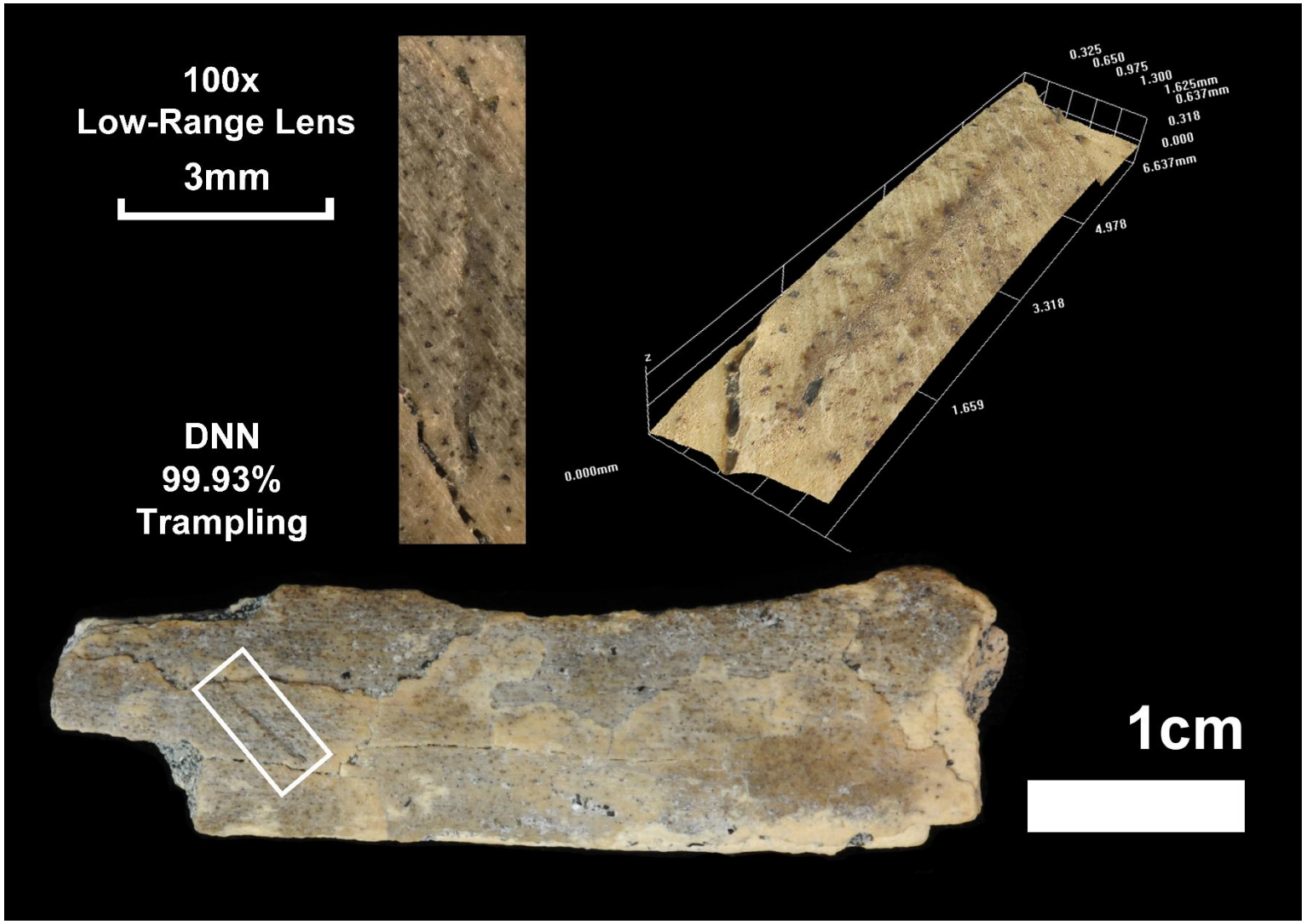


Figure 5.49. – Documentation using 3D Digital Microscopy of taphonomic traces found on an indeterminable bone from the FLK-West archaeological assemblage alongside their classified labels.

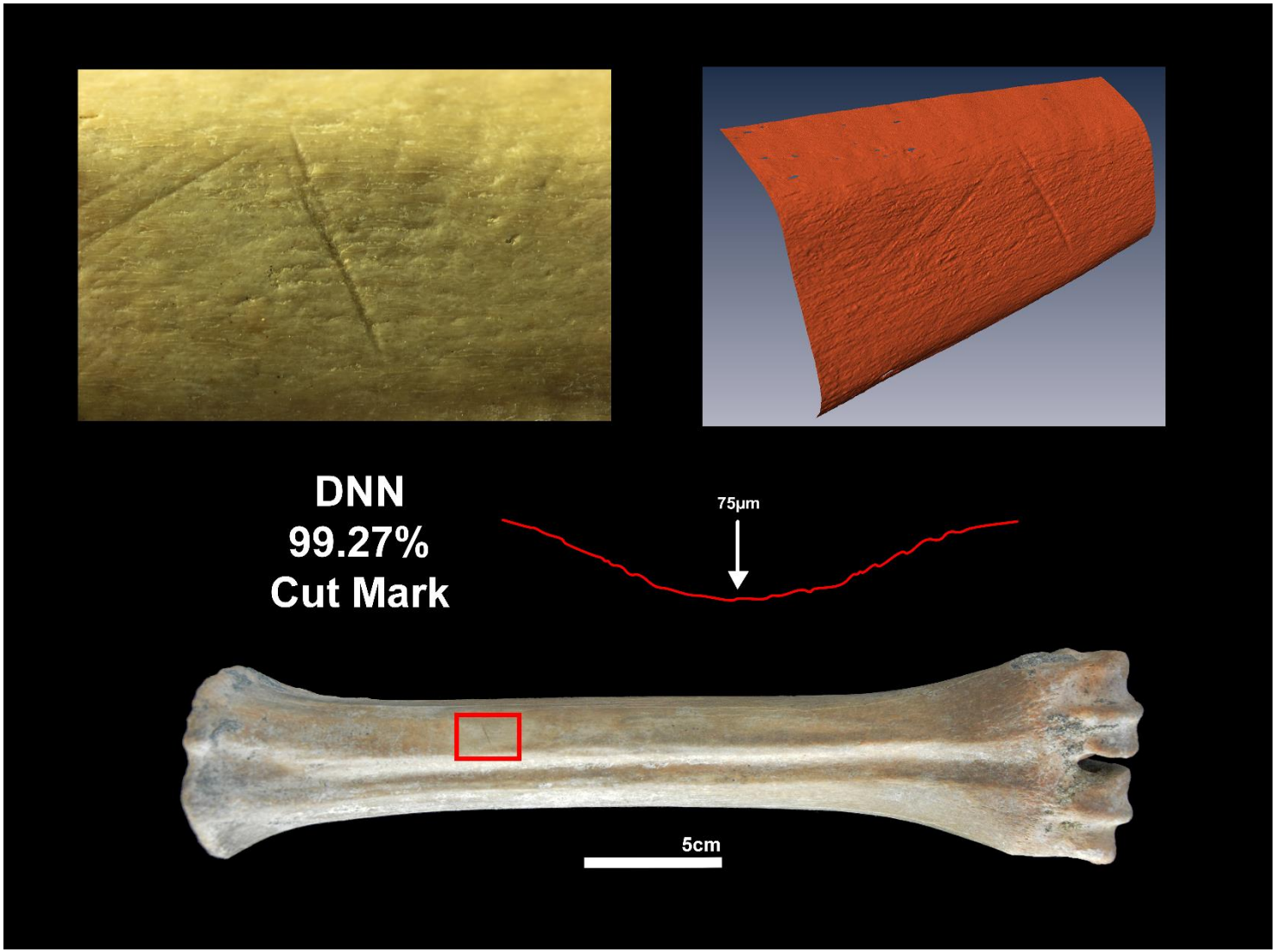


Figure 5.50. – Documentation using 3D Structured Light Scanning of taphonomic traces found on the diaphysis of a Sivatherium's metatarsus from the FLK-West archaeological assemblage alongside their classified labels.

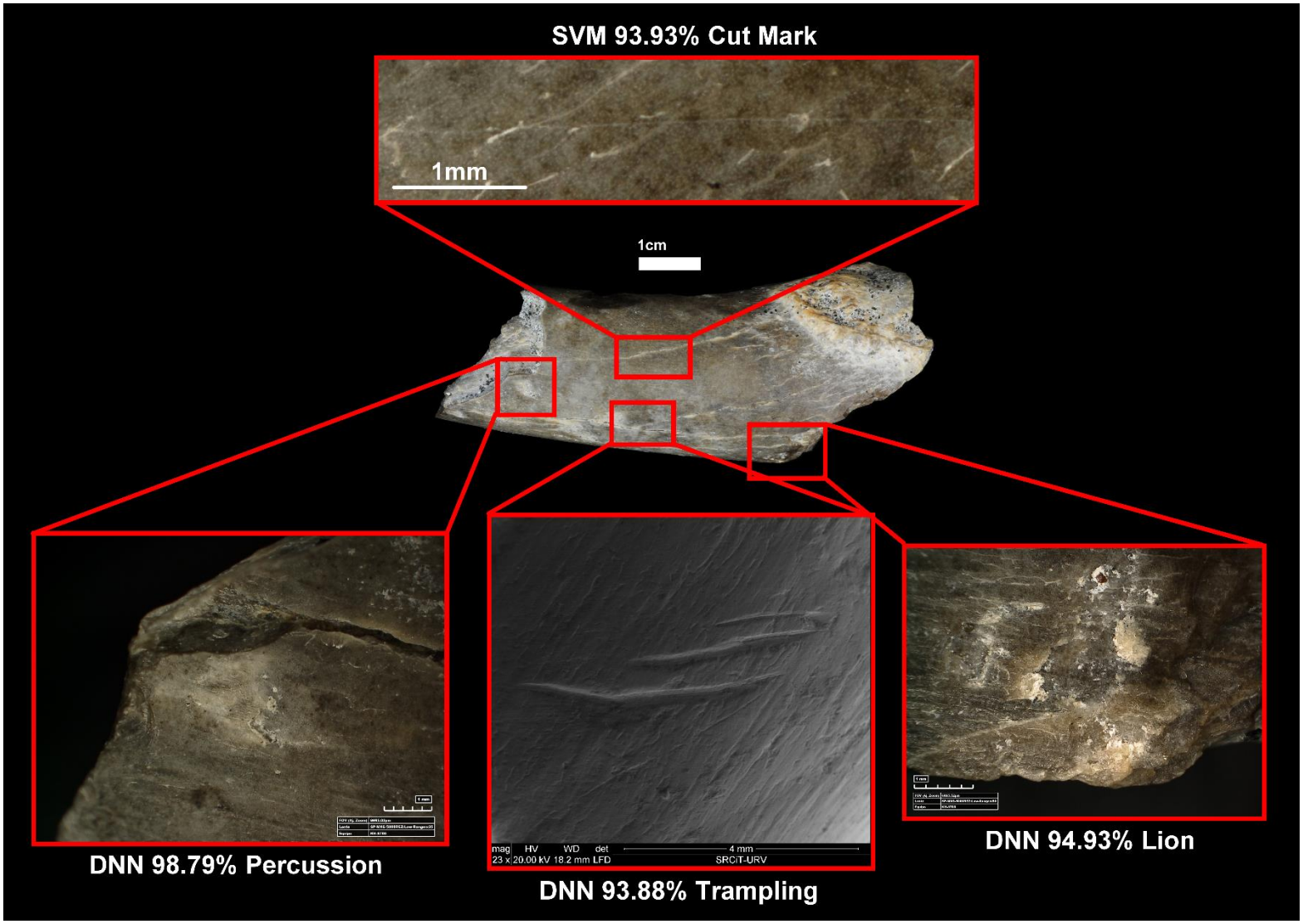


Figure 5.51. – Documentation using ESEM and 3D Digital Microscopy of taphonomic traces found on the diaphysis of a bovid femur from the FLK-West archaeological assemblage alongside their classified labels.

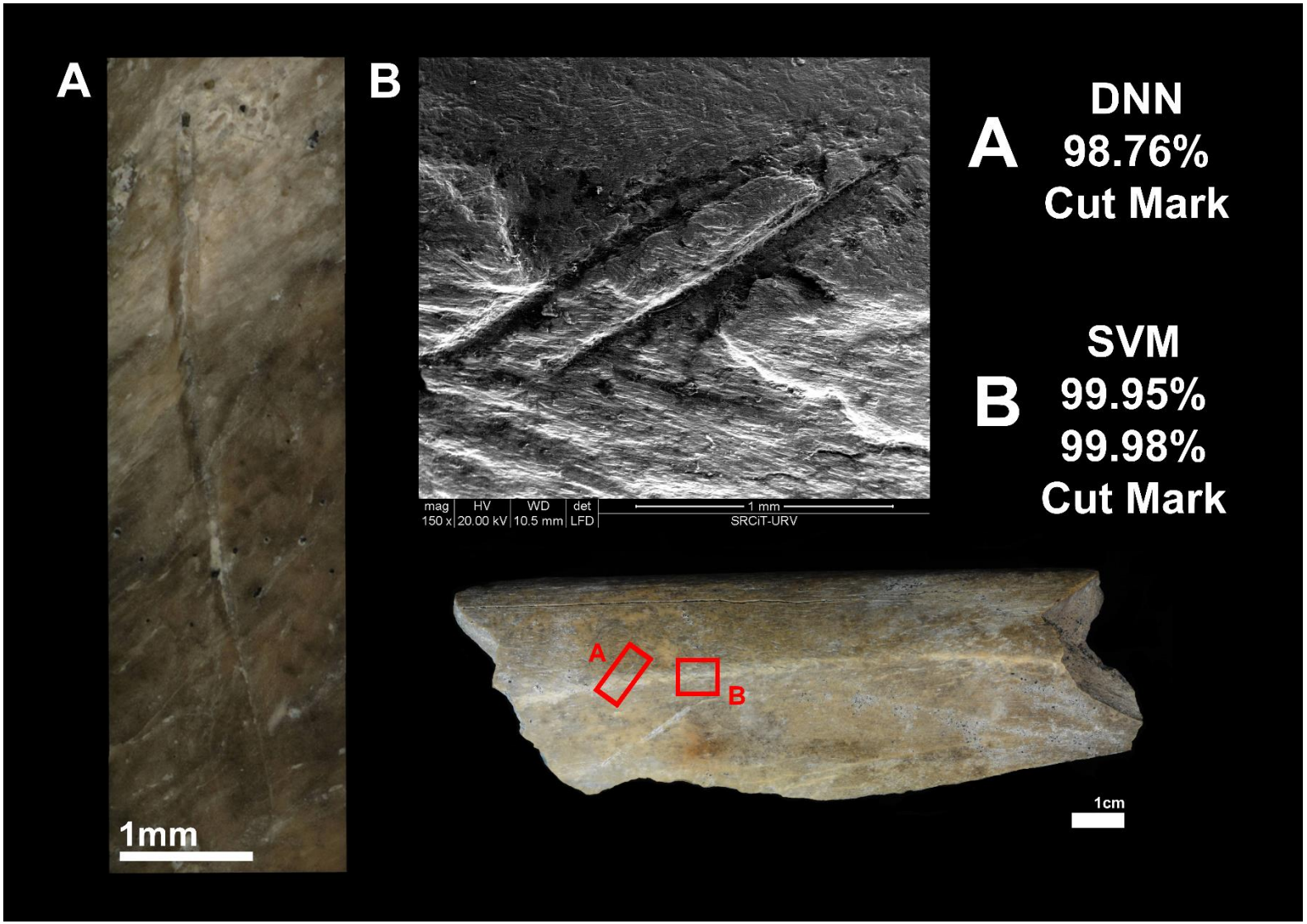


Figure 5.52. – Documentation using ESEM and 3D Digital Microscopy of taphonomic traces found on the diaphysis of a bovid tibia from the FLK-West archaeological assemblage alongside their classified labels.

Chapter 6

Discussion

Nine out of ten? If you were an airline, and you landed nine out of ten planes, you wouldn't exactly brag about that in your commercial.

- Bojack Horseman -

6.1. Microscopy and the Analysis of Superficial Marks

Taphonomic equifinality can be considered one of the greatest issues analysts have to face when considering the interpretation of site functionality and formation. Through the similarities present between multiple agents, as well as the overlapping of multiple destructive processes, finding a means of objectively analysing these processes has become a key component of recent taphonomic research. This line of investigation is especially relevant in sites of great importance for the study of the first hominin populations in East Africa.

This study presents a large quantity of different results that attempt to confront this ever-growing issue, describing new means of quantitatively processing microscopic alterations to bone cortical surfaces. The objectives of this study have focused greatly on the use of advanced statistical methods as well as high resolution analytical tools to process these BSMs. Through this, the current study has been able to present a significant methodological advance in overcoming taphonomic issues in equifinality, and through a preliminary study of microscopic traces on bones from the Early Pleistocene Acheulean site of FLK-West the results have been able to highlight the potential of this particular approach.

Beginning this discussion with the interpretation of the direct results at hand, the first key component of this study addresses microscopy as a means of providing a detailed evaluation of the exact nature of BSMs within a site.

The use of microscopy within archaeology can be considered a fundamental tool when reflecting on the amount of information that can be obtained using this type of equipment. This is especially relevant in taphonomy, when reflecting on how minute alterations to bone cortical surfaces can reveal a series of clues as to the life history of a fossil. Recent advances in the processing of microscopic features have struggled to overcome issues presented when analysing superficial BSMs. In this study, the use of digital microscopy has proven a valuable tool capable of overcoming this. While this approach may seem costly, considering the price as well as the time taken to analyse BSMs, the quality of this tool from a visual analytical perspective is outstanding. Additionally, when considering problematic traces, or sites that may produce a significant paradigm shift to our understanding of human evolution, these disadvantages to this approach should not be considered a discouraging factor. Resolution and scientific objectivity comes at a cost, yet when new data is of a significant value, this value should overcome our adversities and indispositions, favouring science over comfort.

Low cost techniques such as the use of structured light surface scanners have proved successful when studying tridimensional features of taphonomic traces (Maté-González, 2017a, c; Courtenay et al., 2017, 2018b), however fail to provide results when confronting inconspicuous marks. Nevertheless, an important issue to consider is the availability of these pieces of equipment. An additional objective should therefore be to provide comparisons of techniques in order to present a wide range of possibilities that analysts can use to confront different situations, depending on the resources they may have at hand.

Here, the HIROX KH-8700 was successfully able to reconstruct all taphonomic

traces, including carnivore tooth marks and superficial taphonomic traces. In depth assessment of reconstruction quality and variations produced by lighting conditions have been able to refine a 3D reconstruction protocol for these superficial traces. When comparing the HIROX with other microscopic techniques such as SEM, the built in mosaic tiling function is a considerably powerful tool that efficiently creates 3D models. SEMs require additionaly steps in order to compile their images (Vergès and Morales, 2014), making analysis of large surface areas timely and difficult. While 3D reconstruction techniques are indeed available for SEM imagery (Eulitz and Reiss, 2015; Tafti et al., 2015), further investigation is needed in order to assess the efficiency of these techniques and understand their limitations.

3D Microscopy and the use of equipment such as the HIROX has had an increase in use over the last years, with a specific focus on the quality of images and the ability to extract cross-section data derived from 3D images (Blasco et al., 2016; Moretti et al., 2015; Fuentes-Sánchez et al., 2017; Rufa et al., 2017; Duches et al., 2018; Stinnesbeck et al., 2018; Rodríguez-Hidalgo et al., 2018, Under Review). Likewise, confocal microscopy is currently considered one of the most powerful tools on the market, presenting similar options yet at a higher resolution (Archer and Braun, 2013; Maté-González et al., 2017a; Otárolla-Castillo et al., 2017; Pante et al., 2017). Notable publications using this particular approach have focused their efforts on the calculation of volumetric properties of different alterations (Orlikoff et al., 2017; Pante et al., 2017a). Some of these studies have even provided interesting insights into how fluvial abrasion can reduce the volume of a BSM (Gümrükçu, 2017; Gümrükçu and Pante, 2018). Nevertheless, these studies have much to be desired experimentally, where the traces being compared show little taphonomic equifinality to begin with (Pante et al., *ibid*; Muttart, 2017; Keevil, 2018). It would be interesting to see how well these approaches perform when comparing traces that are problematic from a theoretical and practical perspective.

Important comparative studies into the advantages and disadvantages of different techniques have been performed (Maté-González, 2017a, c, Courtenay et al., 2018b), however, the expansion of this research to include more pieces of equipment could be of significant value to the scientific community.

In other instances, combined usage of microscopic techniques have provided interesting insights into multiple aspects of the taphonomic register. Providing exploratory methods for the detection of anomalies on bone cortical surfaces, as well as alterations to the internal composition of organic materials (Bello and Galway-Witham, 2019). Needless to say, many studies are beginning to push towards more quantitative approaches with the aid of powerful microscopic techniques.

The results presented in this study in particular contribute well to these advances, showing how developed digital microscopy can process minute variations in BSMs. Moreover, the detailed analysis of naturally produced linear incisions were able to allow for pattern recognition and the identification of within-group variation amongst trampling marks, thus discovering two different types of traces that have since been labelled *scratch* and *graze* marks (Courtenay et al., 2018b).

While a more traditional mindset into the theory behind trampling mark formation would argue geological and sedimentological factors to be a conditioning variable (Schiffer, 1987; Olsen and Shipman, 1988; Fisher, 1995; Marín-Monfort et al., 2013; Reynard, 2014), in depth analysis of the trampling phenomena in the present study have been able to argue otherwise. This Master's Thesis have been able to reveal other features to be of importance, highlighting how the state of the bone when buried (dry or fresh) is the primary conditioning factor in trampling mark morphology.

Theoretically, this observation can be described through physics using the concept of *cortical hardness*. While multiple authors have argued carcass size to be a conditioning

variable in linear mark dimensions (Pobiner and Braun, 2005; Merritt, 2012, 2016; Pante et al., 2017a), these authors argue this concept from a standpoint that consider differences in bone density. These observations have since been statistically rejected, finding insignificant differences in cut marks produced on different anatomical elements of different sized animals (Moclán et al., 2018; Maté-González et al., 2019). Product of these results, authors argued the true conditioning factor in BSM morphology to be *bone cortical hardness* (Maté-González et al., 2019).

Reflections made by Wallduck and Bello (2018) were the first to hypothetically touch on this, arguing that the stage of carcass decay was a conditioning factor in the type of BSMs observed. These authors argued how stages of rotting as well as rigor mortis were likely to affect the physiological attributes behind butchery activities, whereby the butcher is required to exert more force in order to cut the meat. While this theory may be plausible, it can be argued that a variable of greater importance would be the physical properties of bone and their interaction with the production of cut marks.

An important osteological component that changes over time as a product of decay is the structural integrity of bone (Walden et al., 2017, Öhman et al., 2012). The process of progressive dehydration produces a significant loss in organic material, affecting the properties of osteological materials (Fyhrie, 2010). This essentially produces a change in the hardness of the cortical (Walden et al., 2017), and is likely to affect the morphology of any alteration that damages the cortical's surface during biostratinomy, sedimentation and diagenesis. If the additional presence of periosteum is considered, the likelihood of sedimentary particles scraping the surface of the bone decreases, and the eventual impact of the particles on the cortical surface once breaking through the periosteum thus changes.

The discovery of morphological differences between trampling marks may have numerous impacts on our understanding of the taphonomic record. First, through this study we now know that the hardness, structural integrity and organic composition of the bone's cortical are able to condition the properties of trampling phenomena. Sedimentary components have additionally been seen to simply increase the number of possible trampling marks that can be produced, but not the type. These results have been confirmed using unsupervised ML techniques, as well as through the use of bidimensional analysis. Moreover, the precise conditions under which they are formed are discernible, considering the frequency of marks produced under different experimental conditions. In as such, it can be seen how damage produced by sediments to bone cortical surfaces on fresh bone produce very thin and long trampling marks now known as *scratches* (Fig. 5.3a., Page 105). With the gradual decay of organic material, such as periosteum, the damage produced shifts towards a much wider and shorter mark, now called a *graze* (Fig. 5.3b., Page 105).

Upon observing these frequencies, it may be possible to provide an indication as to a relative timeline for site formation. Considering how trampling marks are [normally considered] diagenetic in origin (Fernández-López and Fernández-Jalvo, 2002), if a predominance of grazes is observed, this would be indicatory of sedimentation occurring after a certain stage of decay. Nevertheless, this rule presents a series of taphonomic assumptions that are not always met: if an animal dies, and sedimentation occurs quickly, the bones would still be fresh, thus indicating a higher likelihood for the formation of scratches. However, if an abrupt change to the sedimentary environment occurs, and the rate of sedimentation decreases with the bones only covered by a thin layer of sediment, this would leave the now dry bones exposed to trampling processes that would generate a higher degree of grazes. From this same assumption, if the bones were uncovered, moved, and reburied in a different context, the type of trampling would vary.

Thus, the real rule here would imply that in a stable sedimentary environment, where

the rate of sedimentation is constant, the ratio of *scratches:grazes* is likely to indicate the state of the bone during diagenesis. This could be indicative also of the rate of sedimentation in this particular environment. However, if the rate of sedimentation is inhomogenous, or postdepositional alterations occur in this sedimentary environment, then this calculation then becomes much more complicated.

In the case of the latter, it could be argued that it may be possible to detect the overlap of traces which may be indicatory of a certain diagenetic timeline (Blasco and Rossell, 2009), however this is relative and definitely in need of further experimentation. Studies into lithic tool dispersals in stratigraphy have been able to identify a larger correlation of artefact damage produced to tools found closer to the surface, with a significantly lower degree of damage when the artefacts are pushed lower down (Gifford-Gonzalez, 1985). While this is logical, it would be interesting to see how a study similar could be applied to bones and to what extent treading damage affects the type of marks observed in relation to their depth.

Dry bones have additionally been noted to be more susceptible to fragmentation when exposed to trampling processes (Blasco et al., 2008). While other experiments have noticed the compactness of the sediment to be a significant conditioning variable in natural fragmentation (Yravedra, 2005, 2006).

Through these several factors, it can be seen how *scratch*:*graze* ratios could produce interesting information about the time scale of different taphonomic processes. Nevertheless, care has to be taken according to the stratigraphic context where the artefacts are located. This is an important step, however, for further investigation into diagenetic processes.

6.2. Trampling vs Cut Marks

Despite the type of trampling, however, equifinality is still present from a theoretical point of view. Considering that even a thin graze could greatly resemble a cut mark, and a graze could resemble a butchery trace such as scraping marks described by Shipman and Rose (1983) and Shipman et al. (1984), the question still lies in how the two can be distinguished (Figures 6.1 and 6.2).

GMM data presented here used to confront this question has proven highly successful, even when using traditional means of classification. Both CVA and LDA are able to separate groups, especially when the two different types of trampling marks are labelled. While lumping the trampling marks together creates quite a high degree of confusion, the visualisation of mean differences using thin plate splines still highlight minute variances in morphological attributes of these traces that can be used for their eventual classification. As has been argued in the past, trampling marks in comparison with cut marks can be greatly characterised by their superficial nature and irregular trajectory (Domínguez-Rodrigo et al., 2009a). What we additionally state here, is how the general dimensions of different traces are to some extent characteristic of the agent that produces them.

From a purely GMM standpoint, the two different traces are significantly different, both considering their distances in feature space and their multivariate characteristics. Additionally, the degree of separation is over 80% accurate using LDA. Nevertheless, the sensitivity and specificity of this traditional model provides a relatively wide margin of error, producing the moderately low kappa value (0.5).

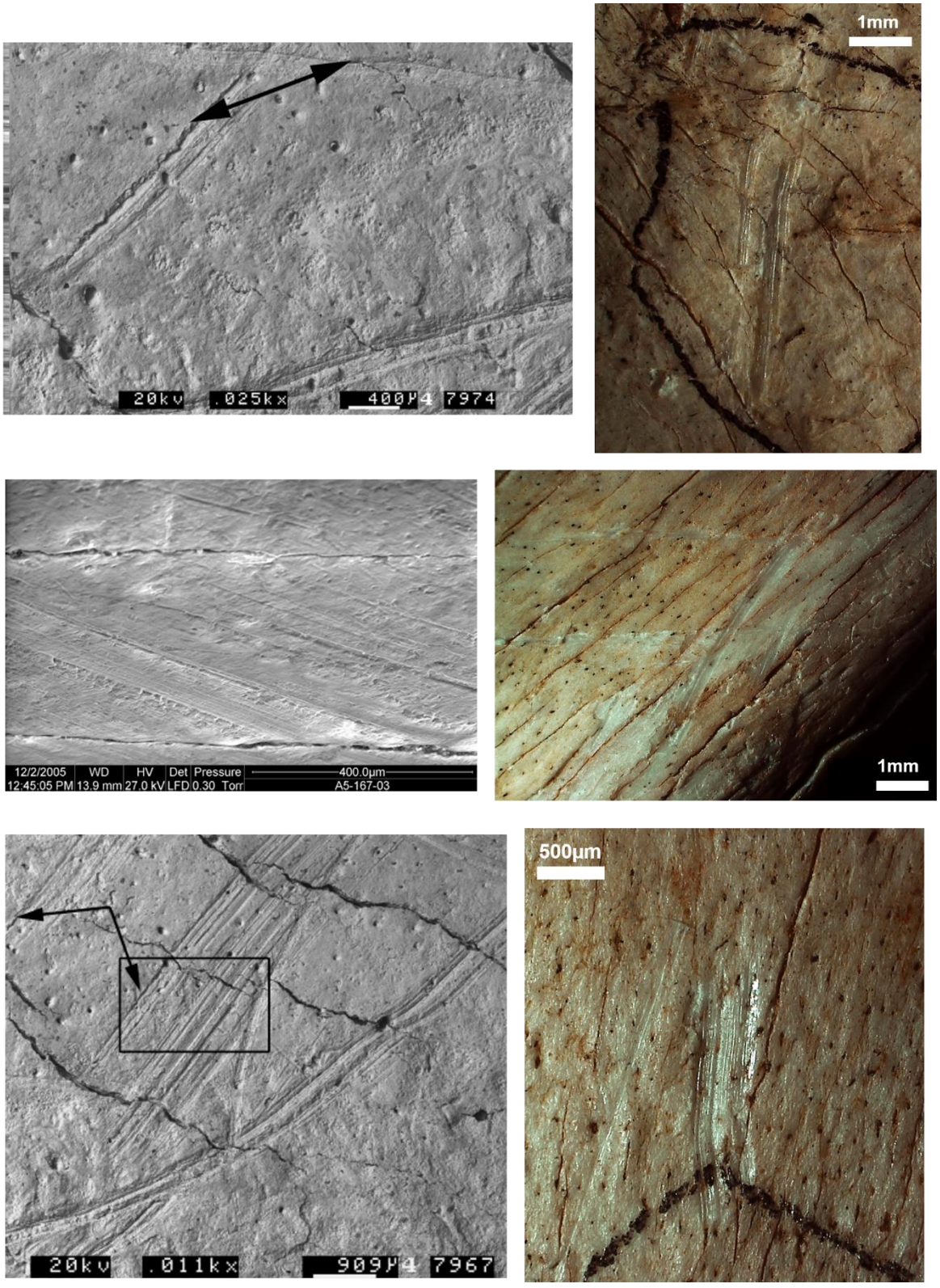


Figure 6.1 – Visual comparison of butchery scrape marks (left) and trampling graze marks (right). Images of butchery marks are from Fernández-Jalvo and Andrews (2016) atlas of taphonomic alterations. Trampling marks were photographed using the HIROX KH-8700 3D Digitial Microscope in the experimental section of this study.

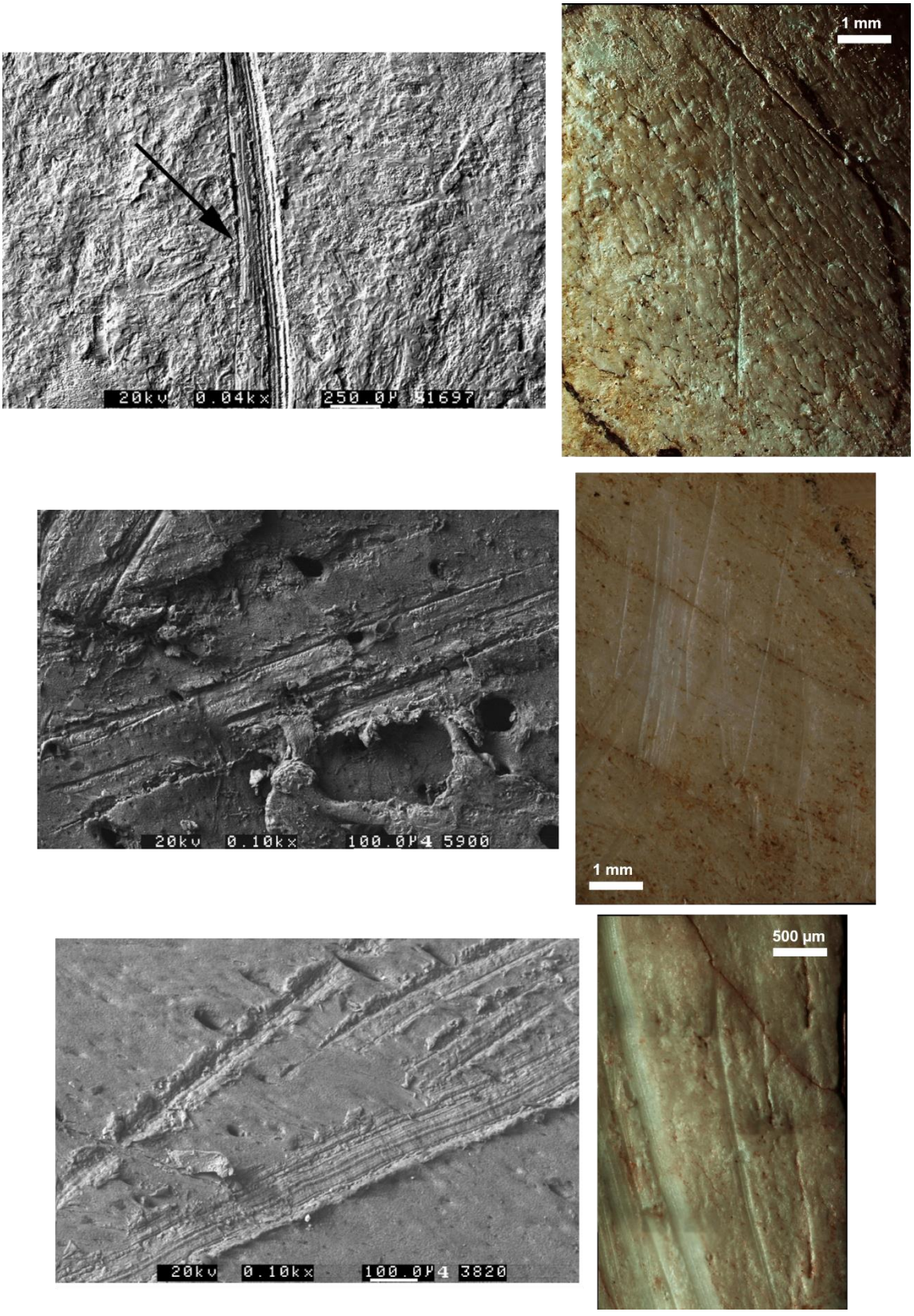


Figure 6.2 – Visual comparison of butchery cut marks (left) and trampling scratch marks (right). Images of butchery marks are from Fernández-Jalvo and Andrews (2016) atlas of taphonomic alterations. Trampling marks were photographed using the HIROX KH-8700 3D Digitial Microscope in the experimental section of this study.

The question is, without the advances of GMMs, would these marks have been distinguishable?

In general, the combined use of advanced microscopy and GMMs were able to provide data that standard taphonomic analyses were unable to reveal. This is especially evident in the case of tooth marks, where classical procedures in the study of tooth mark dimensions are only able to differentiate between different sized carnivores (Sevlaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003; Andrés et al., 2012). While in some Holocene sites these methods are able to reveal data due to the number of possible carnivores interacting are reduced significantly in certain geographic regions (Koungoulos et al., 2018), when considering sites such as those in the Pleistocene, the overlap of many taphonomic agents are very significant (Selvaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003). While 95% confidence intervals are able to reduce the percentage of equifinality to some degree (Andrés et al., 2012), this approach is still not able to clearly divide these samples.

In Pleistocene studies of most of Eurasia and Africa, the presence of hyenas is a significant factor to consider in taphonomic studies (Brain, 1981). Hyenas can produce a very wide range of alterations (Hughes, 1954; Sutcliffe, 1970; Kruuk, 1972; Mills, 1990; Cruz-Uribe, 1991; Lam, 1992; Fosse, 1996, 1997; Brugal et al., 1997; Selvaggio, 1998; Pickering, 2002; Michel, 2005; Diedrich and Žàk, 2006; Ferreti, 2007; Pokines and Peterhans, 2007; Prendergast and Domínguez-Rodrigo, 2008; Lansing et al., 2009; Beauval and Morin, 2010; Kuhn et al., 2010), and when considering the simple qualitative features of a tooth mark, discerning the agent responsible for an accumulation is highly problematic. While some animals such as the crocodile may be distinguishable through the qualitative aspects of their tooth marks (Njau, 2006; Njau and Blumenschine, 2006; Baquedano et al., 2011; Westaway et al., 2011; Njau and Gilbert, 2016; Aramendi et al., 2017a; Domínguez-Rodrigo and Baquedano, 2018), the number of additional cases where this criteria can be used are greatly limited.

When considering the study of linear incisions, the subjectivity produced by each analyst's observations may vary greatly (Domínguez-Rodrigo et al., 2017b), while the criteria established for a study of this nature may not always be applicable (Pineda et al., 2016).

Some specific cases have been able to use these methods to produce significant and interesting results regarding certain sites. In the case of Moclán et al. (2018), cut mark features such as the shape of the groove, presence of shoulder effect, and the nature of internal microstriations were indicative of variations in mark feature morphology produced by specific raw materials present in the site of the Navalmaíllo Rock Shelter. Other cases such as Val et al. (2016), on the other hand, were unable to reach a strong conclusion using this approach. The Noisetier Cave, however, does appear to have a certain degree of post depositional processes which may be altering these results.

The additional results presented within this Master's Thesis confirms the degree of subjectivity and inter-analyst variability, whereby the exact same trampling sample was analysed by multiple analysts and did not produce the same results. This statement should not be confused for an attack at the analysts, but does put in to doubt the value of the method for some of its applications, as proposed by the same author just years later (Domínguez-Rodrigo et al., 2017b).

In the case of FLK-West, for example, the frequency of features noted here have proven incomparable with any of the experimental samples, or any of the published studies presented by Moclán et al. (2018), Pineda et al. (2014) and Val et al. (2016). The presence of water has been significant in alternating the nature of the marks, alienating them qualitatively from any of the compared samples. Hopefully this issue may be overcome by repeating experiments under different

conditions and therefore refining our analogies, however the use of qualitative features will not be the methods used to analyse these final traces.

On all accounts, GMM have proved to be much more powerful for the analysis of these taphonomic traces, especially when used to discern different taphonomic agents.

Considering the results produced by LDA, one of the most commonly used classifiers in modern GMMs, all models were low performing where kappa > 0.8 was unattainable in all cases. In general, the 3D 13-landmark model was the most powerful on all accounts, with Aramendi et al. (2017a)'s model proving to be the most susceptible to misclassification. 2D results obtained from cross sections performed moderately, yet in general, neither of the models could be finally used to efficiently study the archaeological data.

Needless to say, Aramendi et al. (2017a)'s presentation of the 17-landmark model proved to be useful for the differentiation of crocodile tooth marks. These finds can be considered valuable in their initial presentation as they help complement previous work by Baquedano et al. (2012). While traditional approaches are therefore unreliable in some cases, this is not to say they are always poor performers, and the use of other techniques should be dependent on the questions being asked. Nevertheless, in future classification problems, it is highly recommended that old LDA approaches by replaced by more robust ML and DL techniques. The performance of LDA is dependent on the nature of the datasets they are being applied to, therefore, care should be taken in order to provide the most reliable results.

6.3. Artificial Intelligence and Geometric Morphometrics

In this study, AI performance excelled on almost all accounts, with 100% classification of samples being obtained on all samples. The validation of these models on archaeological data proved highly important in the final evaluation of model performance, seen in how C5.0 and RF models were eventually seen to be perfect cases of overfitting.

The overfitting produced by C5.0 and RF models can be explained through the nature of these algorithms. Both techniques form part of a category known as rule based models, whereby the repeated partitioning of data through *if* and *else* statements reaches a final node for classification. Each branches' partitioning of data is produced by the definition of a unique route, which leads to a particular terminal, however in unpredictable cases where the presented inputs can fall in to an infinite number of possibilities, the lack of decision boundaries and confidence intervals most likely causes these models to fail when presented with completely new data. In the case of RF, the advantage of this algorithm and greater predictive capacities can be seen in the bootstrapped bagging method it uses to construct multiple trees, however C5.0's performance relies on a boosting method where a single tree is trained with adjustable weights to fit to the data (Kuhn and Johnson, 2013). While these models have proven highly successful in other case studies (Chao et al., 2014; Richter and Khoshgoftaar, 2018), their adaptability for GMM data is insufficient to be used for GMM taphonomic purposes.

Nevertheless, the use of these models in other applications for archaeology and taphonomy should not be rejected. The majority of these algorithms are adaptable to different types of data, and each should be tested and studied in detail to choose which algorithm works best for the data set at hand. RFs and decision tree based algorithms are highly efficient for sorting categorical variables, thus explaining their efficiency when used by other authors such as Domínguez-Rodrigo (2018)

and Domínguez-Rodrigo and Baquedano (2018). Their partitioning methods also work very well with rounded integers, such as those used by Moclán et al. (2019).

Other poor performing models, such as KNN, NB, MDA and PLSDA can simply be explained through their inner complexity as well as the type of data being analysed. PLSDA and KNN work with the use of Euclidian distances, and thus fail to perform well when the degree of overlapping in the studied sample is high or when the samples being studied display high levels of dimensionality. MDA on the other hand, performs better than LDA because it does not assume linearity, thus overcoming parametric issues presented with some GMM data sets, however in the data sets used here, the complexity of the model is unable to overcome such high overlapping over multiple dimensions. This also explains how MDA works well for processing less complex landmark models. NB, on the other hand, is efficient when used for low-cost computational problems, however the use of Bayesian statistics and probability theories for processing datasets of this type presents too many hidden assumptions that are not applicable on data of this size and type.

The most interesting case of a poor performing model is that of NNET, which has been considered the most popular easily computable technique for data science since the beginning of ML. The failure of superficial neural networks to classify the current datasets, however, is easy to explain. Considering the complexity of the datasets used for training, especially in the case of circular depressions where the number of variables are much higher, it is impossible to efficiently cram 51 variables into a single output. It is logically understandable that a single layer of hidden neurons are unable to produce efficient classifiers.

Why do DNNs and SVMs work so well on GMM data?

Beginning with SVMs, this highly powerful ML and DL model uses hyperplanes and decision boundaries that are constructed in a high-dimensional feature space that helps provide a division between samples. This non-parametric approach is able to transform the feature space, and unlike the decision boundaries in RF, are easily adaptable to new data regardless of their dimensions. The additional ability to deep tune an SVM provides us with one of the fastest, most complex and powerful classifiers available. The ability of SVM to cut through even the densest of feature spaces, such as that of the circular depression GMM form space, proves this model to be one of the most powerful models for this type of data.

DNN on the other hand, unlike SVM and decision tree based methods, is a highly complex mathematical construction, taking a series of values and transforming this data into a final output. The main advantage of this technique, however, is seen in the numerous available hyperparameter optimisation techniques that are available for feature engineering. These models can thus be trained to transform large amounts of data over multiple layers of varying complexities to produce the final results.

6.4. Methodological Reflections

DL is currently considered the future of data science (Domingos, 2015). The efficiency of these models for the processing of very large data sets and production of efficient classifiers, regardless of variable size and complexity, highlights these techniques as the possible future for scientific data analysis.

DL is a field that is growing at a very fast pace. Its introduction into archaeology, however, has been a long time coming. Currently, two main studies include DL; the

first confronts the processing and sequencing of genetic data (Mondal et al., 2018), while the second presents the introduction of Convolutional Neural Networks (CNN) in the processing of taphonomic traces using Computer Vision (CV) (Byeon et al., 2019).

The latter is of great interest to this particular topic, considering how it presents an efficient means of processing photographic images and essentially providing an objective classification of each images' contents. However, it is important to point out some key questions that should be answered before this methodological approach can be applied on a wider scale. As will be discussed, these same issues are especially relevant considering how they are also applicable to this very study.

Byeon et al. (2019) use a highly complex computational algorithm to process each image, pixel by pixel, known as a CNN. CNNs use multiple different computations to perform varying different tasks such as feature extraction (Solem, 2012; Brownlee, 2016b; Bui et al., 2016; Chollet, 2017; Chollet and Allaire, 2017; Zhang et al., 2017; Gu et al., 2018; Spizhevoy, 2018; Zheng and Casari, 2018; Brownlee, 2019b; Zhang et al., 2019). The internal architecture of this particular DNN varies in the additional use of layers that convolute across the image, pool this data into a new array and finally flatten the final transformed tensors into a dense layer typical of the DNNs used in this study (Chollet, 2017). These models require huge amounts of processing data in order to train a final model (Brownlee, 2019b). Many algorithms perform well on the first group of images they are presented with, yet begin to loose performance ability when presented with larger data sets presenting higher degrees of variability. Because of this, most teams working on CV use thousands of images to prepare their models during training.

To provide some examples, the MNIST data set (LeCun et al., 1998a, yann.lecun.com/exdb/mnist/) is one of the most famous open access samples that can be used to train data and consists in a total of 70,000 images, occupying a total of 10Mb for training and 2Mb for testing. These images contain simple numbers from 0 to 10 in black and white, yet are still only able to reach around 99.2% accuracy with approximately 7,000 photos per class (LeCun et al., 1998b). The processing of this sample was additionally performed using an innovative DNN architecture (known as LeNet-5) that has since revolutionised neural network studies (LeCun et al., *ibid*). Another famous data set is the Kaggle Dogs-vs-Cats sample (www.kaggle.com/dogs-vs-cats), which contains 25,000 coloured images that require a series of transformations before they can be used for training. This particular package occupies aprox. 850Mb, requiring large amounts of computational resources and memory for processing. ImageNet (Stanford Vision Lab, 2016, www.image-net.org), the largest CV library and lab in the world, uses 1.3 million photos for multiple label classification algorithms. This particular study also required an intensive investigation into the creation of new DNN architectures (known as AlexNet) in order to process such data (Krizhevsky et al., 2012), and achieves between 62.5 and 83% classification.

While it is true that Byeon et al. (2019) use a standardised process to capture each image, a sample size of 79 photos presents a somewhat optimistic representation of the taphonomic record, especially considering the conditions in which the marks are preserved. Data augmentation may be able to overcome small sample sizes, however for feature detection in CV, large datasets are still needed to provide the model with all the possibilities that could appear in the taphonomic register for the best network weight optimisation and computation.

A multitude of studies exist to discuss the limitations of CV, looking for ways to improve them, considering how light can be considered a highly problematic conditioning factor in CNN performance (Damodarasamy and Raman, 1991), and feature extraction is often difficult in simple structured networks (LeCun et al., 1998b; Krizhevsky et al., 2012;

He et al., 2015; Simonyan and Zisserman, 2015; Szegedy et al., 2015). Alongside this, the possible variability of objects are likely to confuse neural networks, where their primary fault is a lack of *common sense*, a feature that distinguishes human intelligence from artificial intelligence (Thompson, 2018). One example of this *common sense* for most cases of CV can be seen in their perception of perspective (although not necessarily the case here). Through this lack of common sense, the network has to be trained on all the possibilities that it may have to overcome or confront in real-time problem solving opportunities, thus requiring such large data frames for training (Solem, 2012; Redmon et al., 2015; Chollet, 2017; Brownlee, 2019b; Moolayil, 2019). This limitation is fundamental in its robustness in order to confront a constantly changing environment. The complexity of the taphonomic record is absolutely no exception to this.

Additionally, not all bones are easy to photograph, even when using high resolution microscopy, as presented here. The reflective nature of certain surfaces, depending on the diagenetic processes that may have affected the fossil's cortical surface, is a considerably significant factor in the study of BSMs. This issue is also relevant in taphonomic GMMs, whereby the experience of the analyst conditions their efficiency when preparing landmark data. A good photograph or digital reconstruction is not always attainable by some novice investigators. If we were to compare the first images captured of taphonomic traces in this study with the last images captured once the author had gained experience with the HIROX, there is no comparison in quality (Fig. 6.3.). The amount of time spent learning how to use any type of technique, therefore, is a considerable conditioning factor in any type of study.

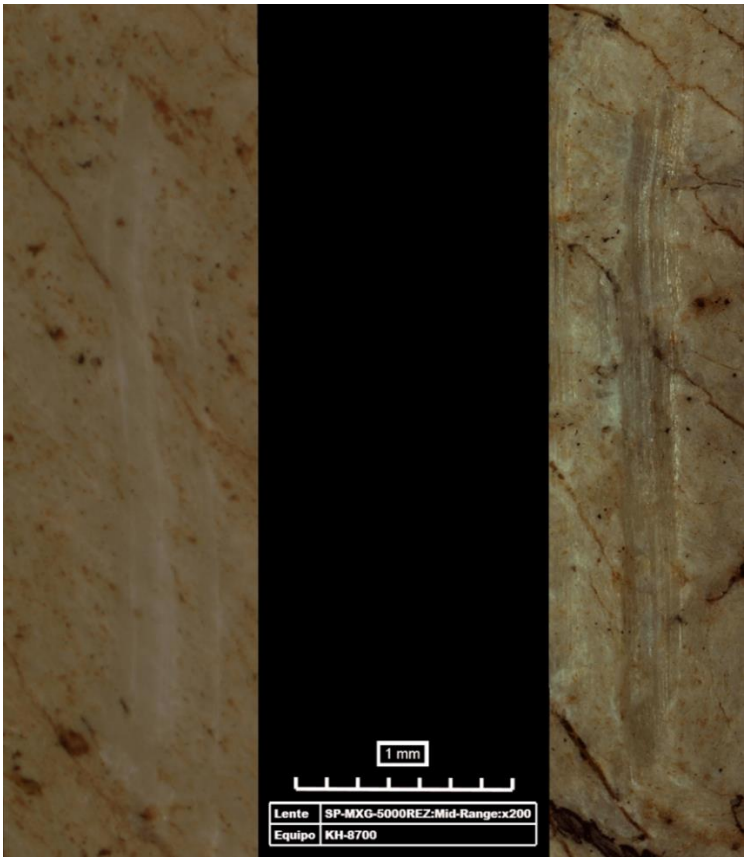


Figure 6.3 – Example of a photo taken of two trampling graze marks when learning how to use the HIROX KH-8700 at the beginning of this study (left) and a photo taken a year and a half later (right).

One powerful technique to overcome some of these issues in CV uses feature extraction and augmentation. These techniques currently occupy a very large proportion of the AI community's research efforts, with the use of DL algorithms such as autoencoders (Chollet, 2017; Géron, 2017; Paterson and Gibson, 2017; Patel, 2019), edge detection algorithms (Wijffels, 2017), and other DNN based regression algorithms (Redmon et al., 2016), to denoise the image and essentially help the process of abstraction on small datasets. This is one way of overcoming some of these issues, however, it would be interesting to see if these algorithms are able to process marks that are significantly affected by other types of taphonomic noise. This noise could include biochemically altered marks as well as incomplete traces, destroyed by fragmentation or fractured specimens.

Nevertheless, it is important to point out that the same limitations are applicable here. A series of disadvantages are still existent for the GMM study of taphonomic traces, even when combined with AI algorithms

It has been seen through the study of FLK-West that equifinality is still present, especially considering the effects of fluvial abrasion. Similar issues have also been observed in other Bed II sites from the Olduvai Gorge, making taphonomic analysis of sites such as TK and SHK much more difficult (Yravedra, personal communication). Some of the traces that have been interpreted as trampling marks present suspicious striations that still greatly resemble cut marks. No other features, however, are present that can confirm this claim. What we have been able to provide here, is an objective means of withdrawing a conclusion, however, we should still consider how true analogy with this site is yet to be reached, and is needed before a full study of the entire faunal assemblage can be performed. Studies into fluvial abrasion have already noted how these types of taphonomic overlap can produce a loss in the volume of traces (Pineda et al., 2014, 2019; Gümrükçu, 2017; Gümrükçu and Pante, 2018), nevertheless, the use of advanced microscopy is a viable option for overcoming this. If we were to employ these high-resolution reconstruction techniques on the correct type of reference sample, SVMs and DNNs are still likely to be able to overcome these issues and refine their classifications. Under the right circumstances, CV is also likely to improve our means of processing this data.

Another issue present, is the completeness of the taphonomic traces being processed. As stated by Gunz et al. (2004) "*Life after death is not kind to [archaeological] specimens*". If we were to consider the scarcity of perfectly preserved traces, we are left with a considerably smaller sample. While the study of trace cross sections is able to overcome this issue, considering how a larger proportion of the mark can be considered representative of the entire mark's morphology (Maté-González et al., 2015, Yravedra et al., 2017c), in 3D models, the loss of a single landmark is enough to completely throw the entire study of a trace.

This has personally been seen on multiple occasions, including the case of some marks in the BK (Yravedra et al., 2017b; Courtenay et al., Under Review), Vallparadís (In Progress) and Cova Foradada (Rodríguez-Hidalgo et al., Under Review) faunal assemblages (Fig. 6.4). All of these sites present typical taphonomic traces that are sometimes presented incomplete through fragmentation, fracturation or the general conservation of the mark with overlapping traces. Thus the number of marks that can be processed using GMMs is limited. This can be frustrating, especially when considering how obtaining a significant sample size in archaeology is often difficult. However, are there means to resolve these issues? The potential of GMM in the study of incomplete specimens could quite possibly overcome this (Gunz et al., 2004a, 2004b, 2005, 2009, 2011, 2012), yet further investigation is needed.

Similarly, do post-excavation factors play a role in these issues? It is common knowledge that taphonomy does not end when the fossil is discovered (Bartram and Marean, 1999; Fernández-López, 2000; Fernández-López and Fernández-Jalvo, 2002). To some,

the role of storage of archaeological materials as well as their treatment in the lab is a significant factor as well (Fernández-Jalvo and Marín-Monfort, 2008; López-Polín, 2012; Krasinski, 2016; Cabec and Toussaint, 2017). In other studies, the restauration of fragile pieces has caused issue when processing certain marks. Consolidation, for example, is a technique employed to increase the structural integrity of the fossil and helps when both extracting fossils in the field as well as preserving them in the storage/museum/the lab (López-Polín, 2012, Appendix 3).

Restauration work in the field, however, is often difficult and hard to perform in the adequate conditions, therefore some bones are often over-consolidated to prioritise saving the piece. While this process is reversible, some still believe that certain restorative treatments may be damaging to the bone (Fernández-Jalvo and Marín-Monfort, 2008). From another perspective, bones that require more drastic techniques for their cleaning may also affect their eventual study. The question would therefore lie in how (or whether) these bones can be safely treated and whether we have a way of overcoming these issues. This is especially relevant in sites where the mechanical cleaning of bones is required for their potential further study in the lab (Valtierra et al., in prep).

Whether or not this issue can be overcome, however, is not necessarily the problem *here*. The point that has to be made, however, is that taphonomists should be aware of any treatment the fossils may have been subjected to over the course of the excavation. Consolidants, for example, can sometimes form a film over the surface of the bone, sometimes making surface examinations difficult (López-Polín, 2012), and frequently reflects light, thus creating distortion in photographs and digital reconstructions (Figures 6.5 to 6.7). This has been seen to produce significant differences (MANOVA $p > 0.001$) in the quality of the reproduction produced, and can thus generates unnecessary noise which eventually hinders their study. Until further investigation is performed confronting this topic, caution is simply advised. While in some cases these errors are easy to detect, in others they may not be as obvious. Most of these problems cause the marks to appear more superficial than they actually are, and would likely cause their misclassification (additional figures and data can be consulted in Appendix 3).

Numerous possibilities exist for potential future investigations into different post-depositional as well as post-excavational processes. Until then, a cautionary note would simply be to consider all the possibilities in a study of any nature. Additionally, it is highly recommended that the conservators work closely with the researchers (López-Polín et al., 2008; Valtierra et al.; In Prep). These limitations in taphonomic contexts show just a small number of the issues both CV and GMM analysts will have to face in the future. One possible means of advancing may include a combined research movement that focuses on the development of DL techniques for archaeological studies, much like those presented by AI communities around the globe (e.g. Anaconda Software Distribution, 2016; Kluyver et al., 2016; Python Software Foundation, 2018; R Core-Team, 2018). These types of collaborations could be crucial for the evolution of our discipline.

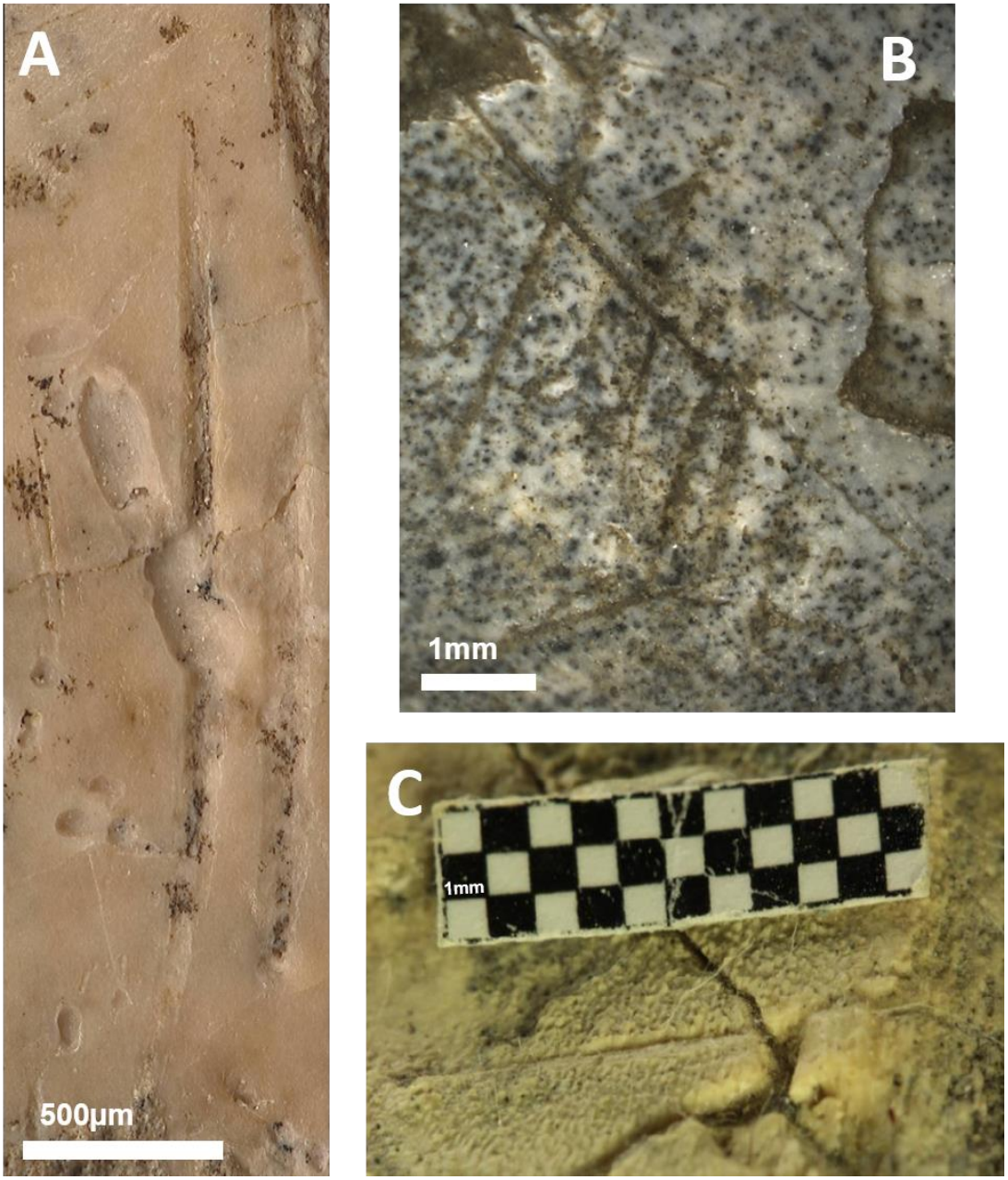


Figure 6.4 – Examples of marks that cannot be processed using the 3D 13-landmark GMM model. (A) A cut mark from Cova Foradada, Cataluña, affected by biochemical destruction to cortical surfaces, erasing landmarks 3, 4 and 5. (B) An unknown linear incision from Valparadís, Cataluña, presenting multiple overlapping striae, sediment within the groove and a fracture to the bone that has destroyed the beginning of the incision, making it hard to locate most landmarks. (C) A cut mark from Bell's Korongo, Olduvai Gorge, Tanzania, where the beginning of the mark has been damaged by a fissure across the bone's surface.

(A) Photo taken by L.A. Courtenay for a study collaborating with Rodríguez-Hidalgo et al. (Under Review).
(B) Photo taken by L.A. Courtenay for a study that is currently in progress. Permission to study sample granted by R. Huguet and I. Cáceres.
(C) *Photo taken by M.Á. Maté-González of a mark that has been studied by L.A. Courtenay (Courtenay, 2017; Courtenay et al., Under Review). Study directed by J. Yravedra*

All marks studied by L.A. Courtenay

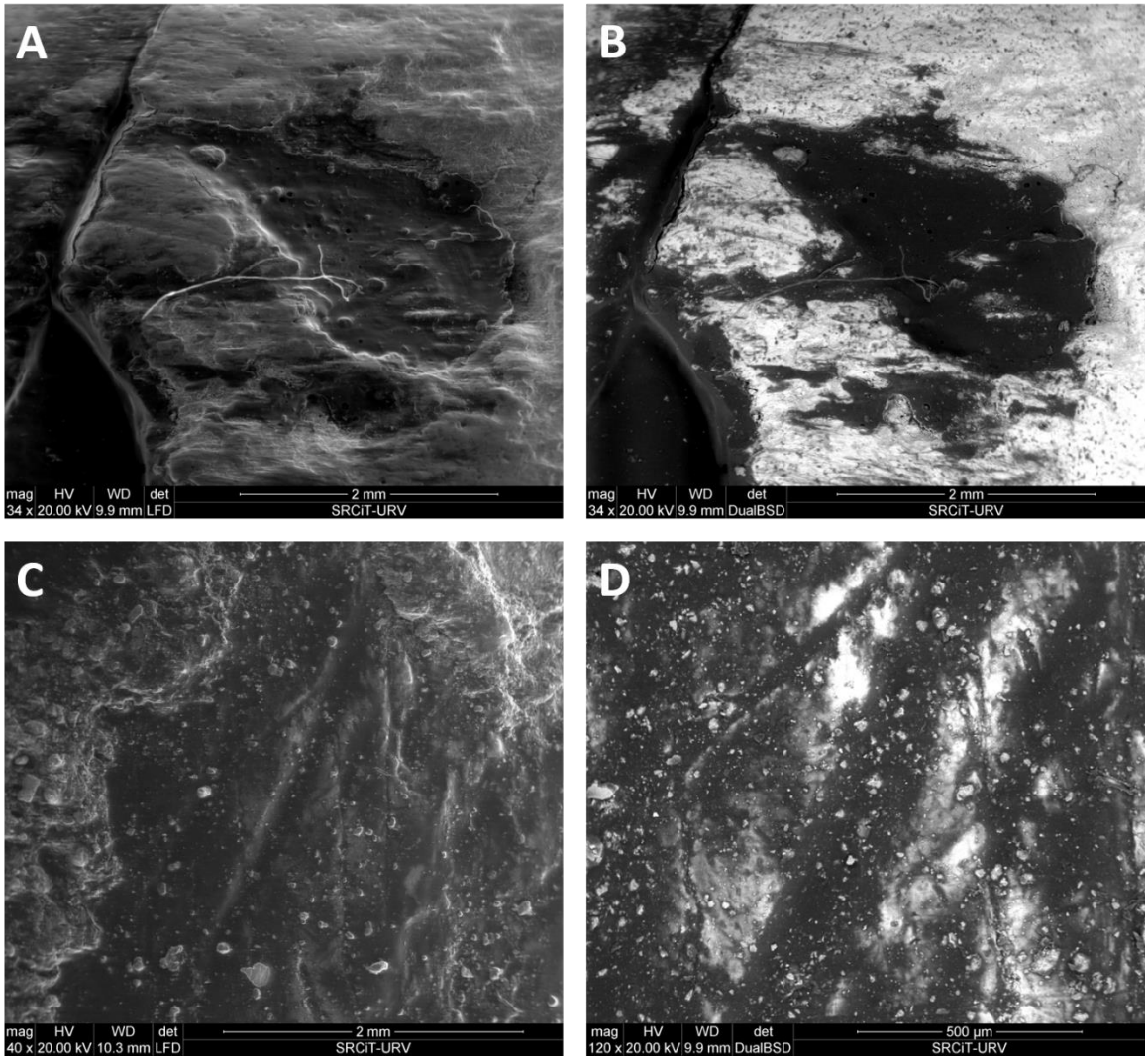


Figure 6.5 – SEM images of circular (A and B) and linear (C and D) marks that have been over-consolidated, requiring further treatment and restauration before they can be fully studied. All marks are from an ongoing investigation into the faunal remains of Valparadís, Cataluña. (A) and (C) are standard SEM images to inspect the topography of the bone. (B) and (D) are obtained using back-scattering electron visualisation techniques. NOTE: in these type of images, darker areas are where the electron beams cannot detect organic material, indicating that the presence of Paraloid B-72 is thick enough to make detection of the bone's surface much more difficult. Chemical data obtained by back-scattering electrons are included in Appendix 3. The consolidation and treatment of these remains in the field have additionally caused particles of sediment, concretions and modern fibres to become glued to the bone's surface.

Images by L.A. Courtenay. Permission to study samples granted by R. Huguet and I. Cáceres.

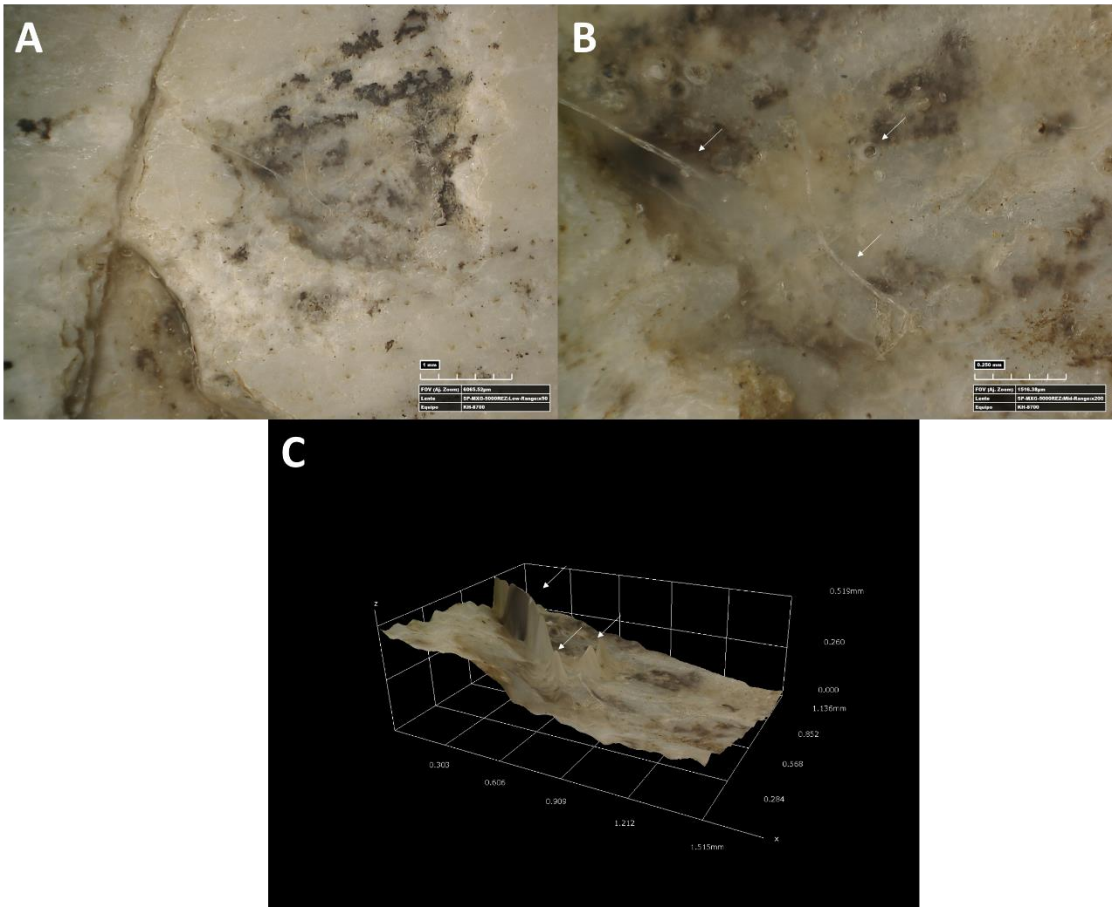


Figure 6.6 – Photos and 3D reconstructions taken using the HIROX KH-8700 of a circular depression filled with a Paraloid B-72 based consolidant. All marks are from an ongoing investigation into the faunal remains of Valparadís, Cataluña. (A) A general photo of the mark and its context. (B) A detailed photo of the mark with white arrows highlighting the presence of fibres and bubbles in the consolidant. (C) A 3D reconstruction of photo B, with white arrows indicating where bubbles and fibres create noise in the reconstruction. NOTE: due to the presence of the consolidant, most landmarks and their corresponding data is unattainable until the consoldiant is removed.

Images by L.A. Courtenay. Permission to study samples granted by R. Huguet and I. Cáceres.

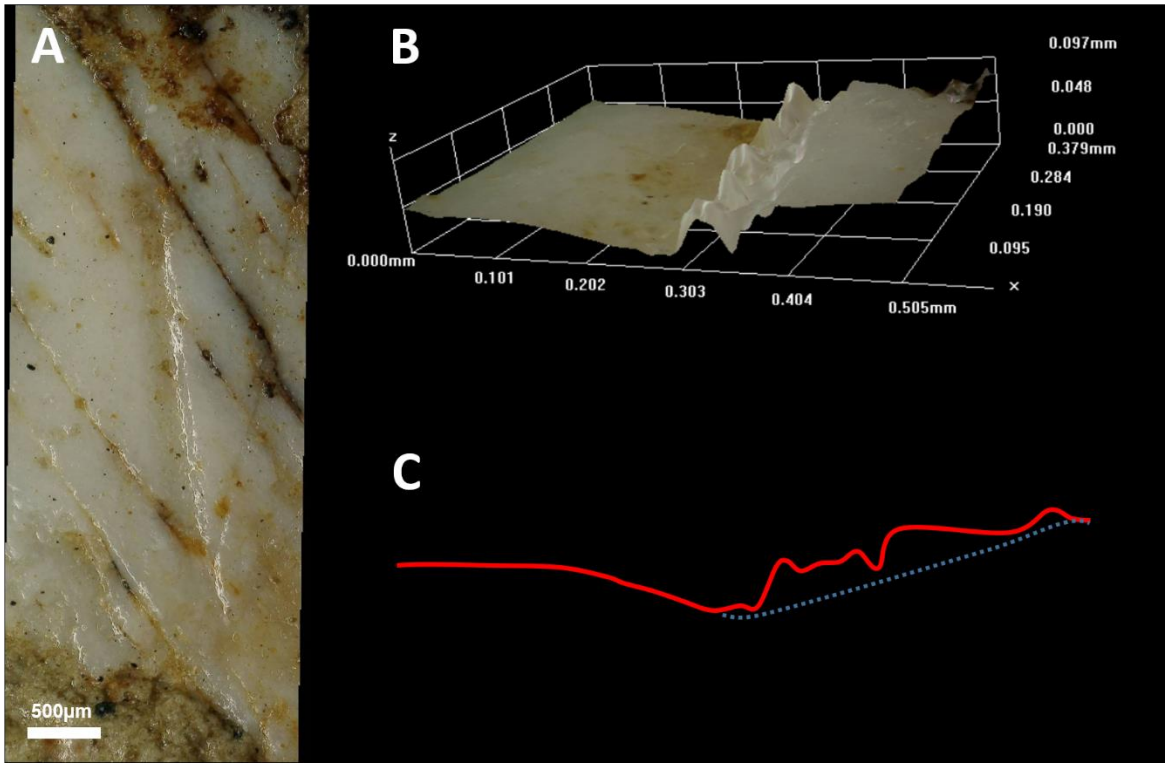


Figure 6.7 – (A) Photo, (C) 3D reconstruction and (C) cross-section profile obtained using the HIROX KH-8700 of a linear incision filled with a Paraloid B-72 based consolidant. Note how the consolidant reflects the HIROX's light, creating clear distortion in the reconstruction of the mark. The Blue dotted line of Image C depicts the predicted real topography of the incision, while the red line depicts the topography detected by the HIROX.

Images by L.A. Courtenay. Permission to study samples granted by R. Huguet and I. Cáceres.

Domínguez-Rodrigo (2018) once stated that the current advances being produced by GMM approaches for cut mark cross sections (Maté-González et al., 2015) are valuable, yet only confront 1 (shape of cross-section) of the 14 to 16 characteristics used in cut mark identification (Domínguez-Rodrigo et al., 2009a; Juana et al., 2010). In fact, if we were to additionally consider how cross-sections also contain information about the symmetry of the groove, this could be extended to include 2 of these features. Courtenay et al. (2017, 2018b, the present study), on the other hand, are able to expand this to 3 characteristics, by including the trajectory of the groove, alongside symmetry and shape which are captured in landmark nº 3.

Nevertheless, the objectives of the aforementioned studies into cross section analysis were originally presented as a means of understanding the raw materials and tools being used to produce these cut marks. No current methods are available using 3D derived 2D data for superficial trace analysis. Considering the practically non-existant nature of most trampling mark cross sections, studies into natural traces thus require a different approach in order to be analysed (Courtenay et al., 2018b).

Moreover, ML and DL allow for algorithms to process data of both a quantitative and a qualitative nature. If GMM data was coupled with other types of datasets, and even combined amongst themselves, these issues could easily be overcome. The conversion of present/absent observations into 1s and 0s is a simple enough task, while the more subjective variables identified could easily be removed. Responding to comments made about the

objective value of some of these variables (Domínguez-Rodrigo et al., 2017; the present study), maybe efforts could be made to try and quantify other characteristics or search for different approaches before combining techniques.

With the introduction of AI techniques in archaeology, some may wonder as to whether ML and DL should begin to replace traditional techniques. The answer here is no. As seen through this study, simple traditional techniques such as bidimensional studies of trampling marks have been included, yielding interesting results. The same way that traditional GMM techniques have updated our knowledge on different types of taphonomic traces. What this study proposes instead, is a hybrid GMM and AI approach to the processing of taphonomic traces, especially for cases where agents prove problematic to interpret otherwise. Standard GMM are needed for the differentiation and in depth analysis of covariation and variation patterns within samples. This is a fundamental step before any type of classification can be performed. Understanding how each variable affects the sample, and finding the necessary theoretical frameworks to truly understand these patterns is fundamental and cannot be performed solely by a computer.

ML and DL instead should be reserved for different tasks such as classification problems or pattern recognition. A mixture of supervised and unsupervised models could provide the perfect toolset to objectively process large data sets.

This study thus presents an exhaustive revision of statistical protocol for GMM studies, of any type. This hybrid Artificially Intelligent Geometric Morphometric approach could thus revolutionise the way data is processed, revealing hidden patterns that we were unable to comprehend or detect before.

It is important to point out, however, that this is not the end to equifinality. This is just the beginning for a more advanced way of attempting to overcome it. Here we present a methodological means of collecting data, with a statistical means of processing it, yet the experimental reference samples could be improved. We are confident that once true analogy has been met, these approaches will still perform greatly when processing the site as a whole.

ML is, contrary to popular belief, actually quite easy to perform, despite the reservations some analysts may have. An example of this can be seen in packages such as 'caret' for R (Kuhn, 2018), which provides a nice toolbox of algorithms that are easily customisable, performing complex ML tasks in a single line of code and providing up to 238 different trainable algorithms. Using caret a simple SVM model can be constructed simply through `train(data, method = "svmRadial")`, which takes milliseconds to fit on any dataset. Furthermore, these models can be saved, shared and reused, making the most of many online repositories and open access publishing platforms (Brownlee, 2016).

While DL is a lot more complex and harder to programme, the results produced by DNNs are incomparable with most traditional analytical techniques. Here was have seen how a combined use of DNN and SVM could efficiently provide classifications of all archaeological taphonomic traces with an average of < 90% confidence.

The sheer size of the AI development community is overwhelming and hard to imagine at times. In the majority of cases, simple google searches for debugging problems and errors produced when doing advanced statistics are able to reveal pages where other developers and programmers have had similar problems. These communities are there to help and make the production of computational applications much easier. This is a fundamental concept that is becoming very popular within science.

How does this system work, though?

An important component of cognition, which links well with studies in Human Evolution, is how information is transmitted. How do we learn? The acquisition of knowledge by a single individual, and sharing

of this information through scientific publications, provides a means of transmitting our different skills amongst a community, thus avoiding the need for each individual to find out these things for themselves. This is a very productive means of evolving as a community, and means that analysts that are just beginning in their academic career can just pick up from where their predecessors left off. In this way, we can focus our efforts on advancing rather than trying to relearn everything that came before us. In this sense, the AI community is a very important practitioner in this movement, sharing all of the information that is being developed for free even down to the code that is being used (e.g. Loizou et al., 2019).

The AI revolution, however, has also sparked off a concept known as the fear of AI (Bostrom, 2014; Domingos, 2015; Tegmark, 2017), which consists in several components (Domingos, 2015; Lanier, 2018). Some feel that computer sciences will begin to take over people's jobs, while others carry a relatively apocalyptic mind-set feel that these advances will teach computers "too much" (which is relatively ridiculous yet more common than one would think). In science a component that is becoming increasingly more evident is a fear found in analysts not wanting to have to learn AI. Many analysts (most archaeologists included) feel that AI requires immense computational power and massive data sets, or requires highly complex programming skills and a developed knowledge of mathematics. This is not necessarily the case. While ML and DL contain a very complex mathematical founding structure, the majority of these mathematical concepts have been translated into simple lines of programming syntax. With the development of packages such as 'caret', for ML (Kuhn, 2018), and 'Keras' (Chollet, 2015), for DL, developers are diluting the huge amounts of mathematical calculations that once formed part of AI and turning them into simple lines of code. With just a simple `train()` or `dense()` function, an investigator with limited knowledge of either mathematics or programming could be capable of developing some of the most complex calculations available for data science (Chollet, 2017). Neither does this actually require that much processing power, as seen through the results presented in this study, the most costly computational model are DNNs, yet most other ML applications only use approximately 30% of the CPU.

What can be seen here is that people interested in learning DL no longer need to spend hours trying to learn lines and lines of java code, yet simply need to open a book or look online to find a simple function that will do all the thinking for them. This is incredibly important for other sciences to evolve on a much broader scale. This has already been seen to work incredibly well for fields of medicine (Webb and Agar, 1992; Mangasarian et al., 1995; Bellazzi and Zupan, 2008; Tosun et al., 2009; Kim et al., 2012; Heredia et al., 2015; Fernandes et al., 2018; Koopman et al., 2018; Richter and Khoshgoftaar, 2018; Sherafatian, 2018; Tapak et al., 2018), pharmaceutics (Ekins, 2016; Gawehn et al., 2016; Kortocov et al., 2017; Han et al., 2018; Mamoshina et al., 2018; Yang et al., 2018) and even business studies (Hegazy et al., 2013; Gomez-Uribe and Hunt, 2015; Böhmer and Rinderle-Ma, 2016; Henrique et al., 2018; Lei, 2018; Noelle, 2018; Bisoi et al., 2019; Long et al., 2019).

ML and DL provides ways of reaching 100% classification, a concept which would have been considered unthinkable only half a century ago (and even less). A valuable quote included at the beginning of this chapter states "*nine out of ten? If you were an airline… you wouldn't exactly brag about that in your commercial."* (Bojack Horseman, S03E06, 03:05). If we were to apply this to a different case study, such as in medicine, when a doctor has an 80% chance of not killing a patient, this is still quite a dangerous number to be playing with. Specialists have therefore pushed towards the development of more secure ways to perform statistical studies, and all of this data is completely open to the public and free.

The revolution of AI is nothing to fear, yet quite the contrary. It provides a new means

of processing mass quantities of data thus liberating analysts to dedicate their time to other things. Archaeology has spent a long time relying on subjective techniques, if we were to update our toolkit, we might be able to confront even bigger questions at a much faster rate. While maths is not the greatest *forté* of most academics in humanities, current research has provided a means of developing algorithms using interdisciplinary teams to help overcome previous boundaries requiring extensive knowledge in other disciplines.

This is a promising time for archaeology as well as taphonomy, considering the advantages advanced statistical approaches have for the field. If archaeology wants to evolve as a *scientific* field, it *must* update its methods, especially in light of recent debates.

6.5. Taphonomic Reflections

The publication of the 'oldest' anthropic evidence of any type is always a problematic issue, usually drawing attention, criticism and eventual debate on the quality of these finds from the entire archaeological community. In African sites, perfect examples of such debates can be observed in claims of the oldest cut marks from Dikika, Ethiopia, (McPherron et al., 2010; McPherron et al., 2011; Thompson et al., 2015; Thompson et al., 2019), which have since been heavily criticised and rejected (Domínguez-Rodrigo et al., 2011; Domínguez-Rodrigo et al., 2012b; Domínguez-Rodrigo and Alcalá, 2016). The current consensus for oldest cut marks in Africa, however, remain to be those of Gona (Dóminguez-Rodrigo et al., 2005), with 2.58 to 2.1 Ma of age, while other promising results have been localised with 1.9 and 2.4 Ma archaeological data from Northern Africa (Sahnouni et al., 2019).

Taphonomic debates revolving around these topics are essential in understanding features of human evolution, considering how current theories argue meat eating to be a fundamental component of our evolution (Washburn, 1963; Lee and DeVore, 1968; Isaac, 1971, 1978, 1983; Hayden, 1981; Bunn, 1981, 1982a, b; Potts, 1982, 1988; Milton, 1987; Speth, 1989; Leonard and Roberston, 1992, 1994; Aiello and Wheeler, 1995; Stanford and Bunn. 2001; Eaton et al., 2002; Ungar et al., 2006). The concept of butchery contains a multitude of different implications, beginning with resource acquisition (Isaac, 1978, 1983; Blumenschine, 1986, 1995; Bunn and Edzo, 1993; Bunn, 1995, 1996; Domínguez-Rodrigo, 1997, 1999; Domínguez-Rodrigo and Piqueras, 2003; Domínguez-Rodrigo and Pickering, 2003; Domínguez-Rodrigo and Barba, 2006; Blumenschine, 2007; Domínguez-Rodrigo and Barba, 2007; Domínguez-Rodrigo et al., 2007), as well as the cognitive technical capacities to manufacture the instruments used for such activities (Schick and Toth, 1998, 2006; Semaw et al., 2003; Toth and Schick, 2009a, b, 2018). Dates of cut marks at 3.3 Ma implicate *Australopithecine* populations to be the first users of tools and butcherers in hominin history, however authors are yet to come to an agreement as to whether these individuals were physically capable of such practices (Schick and Toth, 2006; Key and Dunmore, 2015; Harmand et al., 2015; Domalain et al., 2017). While an argument has been proposed to say that natural edges of unknapped stones could be used for butchery (McPherron et al., 2010), other authors argue that experimentation is yet to be found that supports this claim (Domínguez-Rodrigo et al., 2012b).

Nevertheless, if these finds were to be real, then strong empirical evidence would be needed in support of such a hypothesis. The current study presents a possible methodological approach that could be used to overcome these practical issues. GMMs and AI are efficient and a mostly objective means of

withdrawing very strong conclusions which could be used to pinpoint the precise agent responsible for a faunal accumulation. Given that the correct experimental conditions are employed for comparison. Through AI, new techniques and analytical tools could provide rigid evidence to support an argument, which would additionally then require anyone willing to argue this claim to provide just as valid an empirical counteragument, in order to contribute to the debate.

From a similar perspective, techniques such as these have refined the methodological approaches available for the study of paleoecology, allowing taphonomists to specify precisely the type of agent interacting with the faunal remains. If we were to mix factors of skeletal representation patterns, taphotypes and the morphology of tooth marks, it could be possible to explore in a greater depth than ever before the agents interacting in site formation processes. Faunal remains from sites such as Vallparadís (Martínez et al., 2010; Garcia et al., 2011), which have stirred up a certain degree of controversy (Madurell-Malaperia et al., 2012; Kolfschoten, 2017), could be studied in detail with a means of objectively concluding as to whether the fauna from this site are product of anthropic activity or not.

On a different line of research, the degree of hominin/carnivore interactions can also be identified and studied in greater detail. Throughout the Pleistocene multiple accounts of carnivores hunting hominins have been noted (Brain et al., 1981; Andrews and Fernandez-Jalvo, 1997; Boaz et al., 2004; Yravedra, 2010; Pickering et al., 2011; Domínguez-Rodrigo et al., 2013a; Aramendi et al., 2017a; Deujeard et al., 2017), with an interesting twist to the story when the predator becomes the prey (Münzel and Conard, 2004; Blasco et al., 2010; Rodríguez-Hidalgo et al., 2011; Kitagawa, 2012; Abrams et al., 2013; Gabucio et al., 2014; Camarós et al., 2016). Data of this type is of great interest to understanding paleoecology.

The vast quantity of studies into carnivore paleoecology and development have provided a multitude of different investigatory questions. From one point, the social capacity of Hyenidae populations has led authors to consider the nature of their feeding patterns in the paleontological register (Vinuesa et al., 2015). Similarly, studies into the evolution of the Hyenidae family have come to explain their capabilities of crunching bone (Ferretti et al., 2007), thus explaining their destructive nature. It could be quite possible to use unsupervised learning algorithms for pattern recognition in order to see if different extinct predator species can be separated and studied in greater detail. This theory could potentially propose a means of detecting and studying *Pachycrocuta* in comparison with *Crocuta* through dental morphological variations alone. Through processing sites and observing the degree to which different agents intervened within an accumulation, we may be able to answer several questions regarding the evolution of these animals alongside the evolution of hominins. This could later be extrapolated for more in depth analyses of ecological niches, and even develop network analyses into palaeoecosystems (Roopnarine, 2009; Lozano et al., 2015).

Recent studies into cut marks have been able to provide sufficient resolution so as to decipher the precise raw materials being used (Maté-González et al., 2016, 2017a; Courtenay et al., 2017; Yravedra, 2017b), in some cases even refined to granular composition (Courtenay et al., Under Review). Other experimentation has revealed tool type to be an additional conditioning factor in morphology (Courtenay et al., 2017; Yravedra et al., 2017a). These advances have been able to greatly develop the initial studies into these multiple topics (Walker and Long, 1977; Spennemann, 1990; Choi and Driwantoro, 2007; West and Louys, 2007; Bello and Soligo, 2008; Bello et al., 2009; Domínguez-Rodrigo et al., 2009a; Juana et al., 2010; Bonney, 2014; Weston et al., 2015; Moclán et al., 2018). General studies into cut marks could greatly benefit from a more powerful statistical approach when processing this data, considering the high degree of overlapping and minute variances in some

cases. In such case studies, questions regarding raw material management as well as other socioeconomic variables could be confronted, finding new means to studying lithic tool production/use which could combine well in interdisciplinary studies with use-wear analysts. The functional components of tools are an interesting factor when linking cultural developments with cognitive and behavioural attributes of early humans (Schick and Toth, 1998, 2006; Toth and Schick, 2009a, b; Key, 2016; Key et al., 2016; Key and Lycett, 2017a, b, c). Through in depth studies of different activities that employ different tools, this provides a new perspective to understand a fundamental component of early hominin day-to-day activities.

Cut mark analysis remains to be a very important component of taphonomic studies. Through distribution and concentration, interpretations can be withdrawn concerning type of butchery activity (Martin, 1909; Gifford-Gonzalez, 1977; Binford, 1981, 1985; Bunn, 1981, 1983b; Lyman, 1987; Capaldo, 1995, 1997, 1998; Nilssen, 2000; Bunn, 2001; Lyman, 2005; Domínguez-Rodrigo and Yravedra, 2008; López-Cisneros et al., 2018), as well as certain cultural components. Cut marks on human remains, for example, present new questions concerning possible funerary practices, cultural movements (Stringer, 1985; Cook, 1986; Ulrich, 1989, 2005; Frayer at al., 2006; Walpoff and Caspari, 2006; Yustos and Yravedra, 2015; Bello et al., 2016), and in some arguable cases; alimentation purposes in order to survive (Defleur, 1999; Fernández-Jalvo et al., 1999; Fernández-Jalvo and Andrews, 2003; Valensi et al., 2012; Yustos and Yravedra, 2015; Bello et al., 2016; Rougier et al., 2016). By understanding the different attributes involved in these practices, new patterns may reveal important behavioural components in these populations, aiding in the creation of a more complete reconstruction of past events.

The addition of new analytical techniques and methodological approaches in taphonomy additionally present new means of approaching symbolic behaviour and prehistoric art (Güth, 2012; Bello et al., 2013; Charlin and Hernández Llosas, 2016; Nelson et al., 2017). Through linear incision analysis, we can find new means of determining whether marks are intentional or not, and better refine the image of artistic engravings. These techniques could additionally develop our understanding of tool functionality such as that of the burin (Moretti et al., 2015), with the additionally possibility of extrapolating new techniques to also analyse patterns in fluted cave engravings such as those from Rouffignac (Van Gelder, 2010; Cooney and Van Gelder, 2011; Van Gelder, 2014, 2015; Van Gelder and Sharpe, 2015).

Here are just a few possibilities these new methodological approaches may present for prehistoric studies. With a refined approach to traditional studies, taphonomists may be able to overcome previous barriers and provide greater resolution in the study of these sites. Additionally considering how taphonomic analyses are not exclusive to the study of BSMs, combined with analysis of skeletal profiles and fracture plane analysis (Arriaza and Domínguez-Rodrigo, 2016; Egeland et al., 2018; Moclan et al., 2019), we are a step further in forming a very efficient means of processing the taphonomic register in greater detail.

One particular cases study where these approaches will be employed are sites from the Olduvai Gorge. The Olduvai Gorge is a key archaeological complex for the study of early hominin populations. Sites in both Beds I and II, especially the case of FLK-Zinj, have been referred to as the "*testing [grounds for] virtually most behavioural models that have been proposed for early Homo in the past 50 years*" (Domínguez-Rodrigo et al., 2019). The present study is an additional case of this.

Zooarchaeological evidence from these iconic sites have presented evidence of exploitation of macro and megafauna as far back as ≈ 2 Ma. The intense trophic pressure seen through intervention of lions, hyenas and hominins in the formation of most faunal assemblages provides interesting data that can be used to reconstruct the paleoecology of some

of the first *Homo sp., Homo habilis, Homo erectus/ergaster* and *Paranthropus boisei* populations.

6.6. Archaeological Reflections

This particular study refers primarily to some of the first Acheulean populations currently known in Africa. The current understanding of the Acheulean technocomplex cites its origin in Eastern Africa. Key sites contributing to this knowledge include Kokiselei, Kenya (Roche et al., 2003; Lepre et al., 2011), Konso, Ethiopia (Beyene et al., 2012), and FLK-West, Tanzania (Diez-Martín et al., 2015), associated with a chronological window of ≈ 1.75 to 1.7 Ma.

The profile of FLK-West, while preliminary, links well with finds from Konso, presenting abundant faunal remains of varying sizes, including *Sivatherium, Elephas, Hippopotamus, Ceratotherium* and *Diceros* alongside many species of *bovids* (Suwa et al., 2003). All faunal remains present frequent traces of butchery as well as abundant carnivore and crocodile teeth marks (Echassoux, 2012). The additional value of this site is greatly seen in the presence of both hominin fossils attributed to the species *Paranthropus boisei* and *Homo erectus*. Kokiselei, on the other hand, does not present a direct association with faunal remains, yet *does* have a valuable lithic assemblage with LCTs that appear similar yet slightly cruder to those found in FLK-West (Chevrier, 2012). The great importance of FLK-West in comparison with these other sites, however, is that this is the first site to have the archaeological materials dated directly (Diez-Martin et al., 2015), and indeed present a greater complexity in the skill required to produce the pieces found here (Sanchez-Yustos et al., 2017b, 2018).

Understanding the diet of this early cultural movement is of growing interest to the archaeological community, especially when trying to find the change from Mode 1 to Mode 2 type lithic technologies. Traditionally, some were led to believe that the appearance of LCTs, such as bifacial handaxes, were attributed to different butchery activities (Clark and Haynes, 1970; Isaac, 1977; Keeley, 1980; Nowell et al., 2016), while others argued them to be multifunctional instruments, much like a prehistoric Swiss army knife (Evans, 1872). Nevertheless, technofunctional studies as well as microscopic analyses of these tools lead to different conclusions, highlighting the working of vegetation (Toth, 1985; McNabb, 1989; Jones, 1994; Domínguez-Rodrigo et al., 2001), and digging (Toth and Schick, 2009a), to be important activities, among other specialised tasks (Key et al., 2016; Key and Lycett, 2017a, b, c). It is also true that the appearance of these more developed tools is not associated with a drastic change in the size of animals being consumed, nor the intensity at which they are being hunted (Domínguez-Rodrigo et al., 2009b, 2013b; Organista et al., 2016, 2017; Organista, 2017). Therefore, the development of these cutting tools may not actually be associated with a change in hunting strategies, yet probably presents a greater correlation with the development of early hominin abilities to manipulate their surroundings in order to survive (*ibid*; Sánchez-Yustos et al., 2017a; Yravedra et al., 2017b; Courtenay et al., Under Review).

Provided that initial preliminary studies of the few cut marks available in FLK-West are yet to show evidence that handaxes were used for butchery (Yravedra et al., 2017a), it would be interesting to see how this study progresses when more bones are obtained from the site over the course of the different levels. An additional interdisciplinary study of technofuntional and potential features of these tools alongside microscopic evidence from both a traceological (lithic) and taphonomic (fauna)

point of view may be able to uncover these details. The approach in the present study would be able to potentially answer these questions, however considering how fluvial abrasion seems to make marks appear wider and shallower, these criteria correspond well with the type of marks that bifacial handaxes produce (Courtenay et al., 2017). This would most likely produce an overclassification of marks as product of bifacial handaxes until more experimentation has been completed. The current line of investigation is thus of great value when considering the advantages these high resolution techniques may produce.

Considering the large size of some of the animals presented within this site, including the classification of cut marks on megafaunal remains such as *Sivatherium* (DNN = 99.27%), we have strong evidence to suggest that humans played an important role in the accumulation of faunal remains from this assemblage, with the additional presence of lions in FLK-West interested in the same resources (DNN = 95.18%).

The trophic pressure between hominins and lions, in this case, is of great interest to the future studies of FLK-West. Current understanding of the paleoecology of Beds I and II highlight an important presence of feline activity in the lower levels of the gorge (Blumenschine, 1995; Capaldo, 1995; Pante et al., 2012; Domínguez-Rodrigo et al., 2007; Arriaza and Domínguez-Rodrigo, 2016; Aramendi et al., 2017b). The levels of Bed II are characterised better by the increased presence of hyenas in most sites (Monahan, 1996a, b; Domínguez-Rodrigo et al., 2007; Egeland and Domínguez-Rodrigo, 2008), nevertheless, considering how FLK-West lies on the very bottom of Bed II and is also associated with the great Olduvai palaeolake, it is possible that the very beginning of the Bed II is very similar to that of the FLK-Zinj palaeolandscape. The identification of lion chewing in the preliminary study here are too few to draw any type of conclusion, yet once again it would be interesting to see how the results develop throughout the following years of excavation.

The competitivity for resources, however, is an interesting line of research that could highlight multiple aspects of early hominin lives during the beginning of the Acheulean. The use of high resolution technologies and advanced statistical analyses could begin to unravel the taphonomic register of this site and help reconstruct the paleoecology and hunting habits of these populations.

Acquisition of resources is still a topic of discussion to this date. When considering how hippopotamus, for example, were an animal to be feared throughout ancient history (Parrinder, 1971; Eliade, 1991; Redford, 2002; Stünkel, 2017), it is hard to imagine that these less developed populations were able to hunt these animals. When confronting an animal such as a hippopotamus or a rhinoceros, there is a great possibility that obtaining these animals could cost the hunter their lives, unless a highly developed hunting strategy was employed. These strategies could consist in communal activities (Binford, 1978; Forbis, 1978; Speth, 1983; Driver, 1990, 1995; Steele and Baker, 1993; Frison, 2004; Meltzer, 2006) – a common practice detected in multiple Middle Pleistocene sites (Lagercrantz, 1934; Jaubert et al., 1990, Brugal, 1995; Gaudzinski, 1995, 1996; Costamagno et al., 2006; Rendu et al., 2009, 2012; Discamps et al., 2011; Rodríguez-Hidalgo, 2015; Rodríguez-Hidalgo et al., 2017; Agam and Barkai, 2018). Strong evidence exists to argue possible complex organised hunting in the upper Bed II of the Olduvai Gorge (Organista et al., 2016; Organista, 2017), therefore it would be interesting to see to what extent this can be inferred here. Furthermore, evidence from BK support that these hunting capacities were even present in Developed Oldowan populations (*ibid*; Sánches-Yustos et al., 2017a), thus arguing that the capacity to hunt is not exclusive to the Acheulean.

Whether the animals in FLK-West were hunted or not, the presence of intense anthropic consumption of these animals has

been confirmed here through objective empirical data. The presence of cut marks and a percussion mark is indicatory of consumption of these animals, regardless of whether access to these carcasses was primary or secondary. While the sample is small, and the activity of some agents may appear insignificative when a larger samples is obtained, the process of breaking the bones to obtain marrow is an important component of understanding human access to animal resources. The consumption of marrow is direct evidence of intensive processing of animal remains to make the most of all components available. In addition to this, when considering the presence of other percussion marks identified in the overall sample (Yravedra and Domínguez-Rodrigo, 2018), the additional presence of bone flakes and a number of cut marks on high meat-bearing anatomical elements of medium to large sized animals implicates that these early Acheulean populations had access to high-quality resources.

Forming part of the same palaeolandscape, the site complex of HWK (including HWK, HWK-E and HWK-EE), presents a case of Oldowan technologies associated with similar degrees of consumption of large animal carcasses and intense presence of carnivores (Pante et al., 2017b; Torre and Mora, 2018). HWK and FLK-West *could* be contemporary, however, at the present moment in time there is no means of confirming this. The only conclusions that can be currently withdrawn associate both sites to the same palaeolandscape (Uribelarrea et al., 2017). While some debate exists as to which hominin populations were responsible for the different accumulations, some argue that the possibility of *Homo habilis*, *Homo erectus/ergaster* and *Paranthropus boisei* all coinhabited this landscape (Pante et al., 2017b). Through these sites, however, it can be seen how Oldowan and Acheulean populations may have coexisted, yet the reasoning behind this is not yet clear. Did some of these populations have different socioeconomic requirements to others? To what point is this movement cultural? What can be made clear, is that this particular paleolandscape is of great interest to the emergence of the Acheulean and may provide interesting data as to its transition.

Both the areas of HWK and FLK-West present significant fluvial abrasion on a lot of the osteological material present (Pante et al., 2017b; Yravedra et al., 2017a; the present study). Their association with the complex fluvial system that ran towards the Olduvai palaeolake raises a number of taphonomic issues that require developed statistical tests in order to resolve. This is a fundamental concept considering the problems that many Acheulean sites seem to face.

Moreover, from an interesting perspective, one of the problematic features of many Acheulean sites remains to be the lack of fauna in association with these tools (Isaac, 1977). While for some sites this is dependent on diagenetic agents (Pappu et al., 2011; Pappu and Akhilesh, 2019), the increased importance of biostratinomic agents such as fluvial alterations are becoming an increasingly important topic of investigation (Domínguez-Rodrigo et al., 2012c, 2014b, 2017a). Studies of this nature tend to confront spatial distributions (Lenoble and Bertran, 2004; Pante and Blumenschine, 2010) and minute variances in orientation (Walter and Trauth, 2013; Cobo-Sánchez et al., 2014). In general, while the remains from FLK-West appear to have a minute variation in spatial distribution (Diez-Martín et al., 2015; Uribelarrea et al., 2017), the degree of fluvial abrasion to cortical surfaces is significant. Considering the results presented here, the predominance of grazes within this preliminary study is indicative of trampling processes occurring while the bones are dry. Combined with the knowledge that none of the bones from the entire sample present weathering, subaerial damage and a small percentage of breakage occurred while the bones were dry (<40%: Yravedra and Domínguez-Rodrigo, 2018), the taphonomic register is enough to suggest that the exposure of these remains to biostratinomic processes was very short.

A combination of these different features with taphonomic data could be of great interest for FLK-West. By examining in detail the taphonomic register of this site, a more in depth analysis could be able to decipher the formation of the site and understand the precise nature of the populations associated. As can be seen throughout this study, the questions that can be proposed and answered by these methodological and statistical approaches are multiple, and as a means of confronting a site such as FLK-West, these questions become increasingly more important.

Before moving on to our conclusions, a final theoretical reflection is to be made.

Fenn and Raskino (2008) propose that with any new demonstration of innovative research, what is triggered is known as the '*Gartner's Hype Cycle*' (Fig. 6.8). This begins with an *innovative trigger*, whereby the demonstration of the new concept is introduced, and through a certain period of time, the developers are able to express the potential of their findings. This initial wave of interest eventually builds potential with the provided excitement of future possibilities until the theory reaches a *peak of inflated expectations*. This is where researchers are keen to apply the technique, and the number of publications using these methods increase greatly. The unfortunate part of the cycle, however, is known as the *trough of disillusionment* and is presented through the eventual discovery of the method's limitations, seen in the loss of popularity of said innovation. In order for the new concept to survive, the *slope of enlightenment* and *plateau of productivity* is to occur, whereby researchers eventually find ways around the limitations and analysts can begin to use these techniques on a broader scale.

The *Hype Cycle* is an interesting theory that has been observed in multiple marketing and business schemes, yet is also applicable to research and investigation.

To provide an example, the techniques proposed by Silvia Bello are of increasing popularity (Bello, 2011; Bello et al., 2013; Moretti et al., 2015; Maté-González et al., 2015, 2016; Arriaza et al., 2017; Fuentes-Sánchez et al., 2017; Courtenay et al., 2018a, b; Duches et al., 2018; Stinnesbeck et al., 2018; Yravedra et al., 2017a, b, c; Bello and Galway-Witham, 2019; Maté-González et al., 2017a, b, 2019), yet their research has indeed been around for longer than most (Bello and Soligo, 2008; Bello et al., 2009). GMMs, on the other hand, is a very popular technique in physical anthropology (as was seen in Chapter 2.7.2 – *Geometric Morphometrics: Applications in Archaeology*), yet the introduction of GMM in taphonomy has only been around for 4 years, and the number of teams employing these techniques are currently limited to a smaller number of analysts (note publications by the team of J. Yravedra, including multiple studies by Maté-González, Courtenay, Aramendi and Arriaza). From a different perspective, work into taphonomic trace analysis using other advanced microscopic approaches are also only just beginning (Orlikoff et al., 2017; Gümrükçu, 2017; Gümrükçu and Pante, 2018; Otárolla-Castillo et al., 2017a, b; Pante et al., 2012, 2017a, b; Muttart, 2017; Keevil, 2018).

AI, on the other hand, is sufficiently new in archaeology that less than 10 articles are currently available using any type of ML or DL techniques (Arriaza and Domínguez-Rodrigo, 2016; Derech et al., 2018; Domínguez-Rodrigo and Baquedano, 2018; Domínguez-Rodrigo, 2018; Egeland et al., 2018; Mondal et al., 2018; Byeon et al., 2019; Courtenay et al., 2019; Moclán et al., 2019).

The specific pattern here can obviously see how these new techniques are on the rise, which would predict that the *innovative* trigger has just begun for some. Care has to be taken, however, to ensure that the peak rises and if sufficient precautions are taken, the *trough of disillusionment* may not even be necessary. Unlike the "hidden mathematical assumptions" that some authors claim to be an issue with statistics (Stapert and Street, 1977), archaeologists are required to update their knowledge on the current procedures available before blindly applying techniques to their data. For ML and DL, this could be as simple as

opening a book (Chollet, 2017; Chollet and Allaire, 2017; Brownlee, 2016a, b, 2019a, b).

Here it is argued that the largest obstacle these approaches currently have to face is the inertia archaeology seems to have at the moment. Analysts may fear that ML and DL is difficult and costly, however, new ways of making these types of study easier are appearing every day. The biggest requirement needed for ML and DL to become an important part of archaeology is that more teams should begin to collaborate and try and update their methods, considering the advantages these approaches may have for many other sites. It would also be of great interest to begin to spread AI algorithms into other disciplines as well, such as lithics, associated use-wear analysis, palaeoanthropolgy (through GMMs or otherwise), dental micro-wear, spatial archaeology, prehistoric art, palaeobotany, geoarchaeology, restauration and even the study of late prehistoric and historic materials. AI, if considered the future of anything, should be considered the future of archaeology, not exclusively taphonomy.

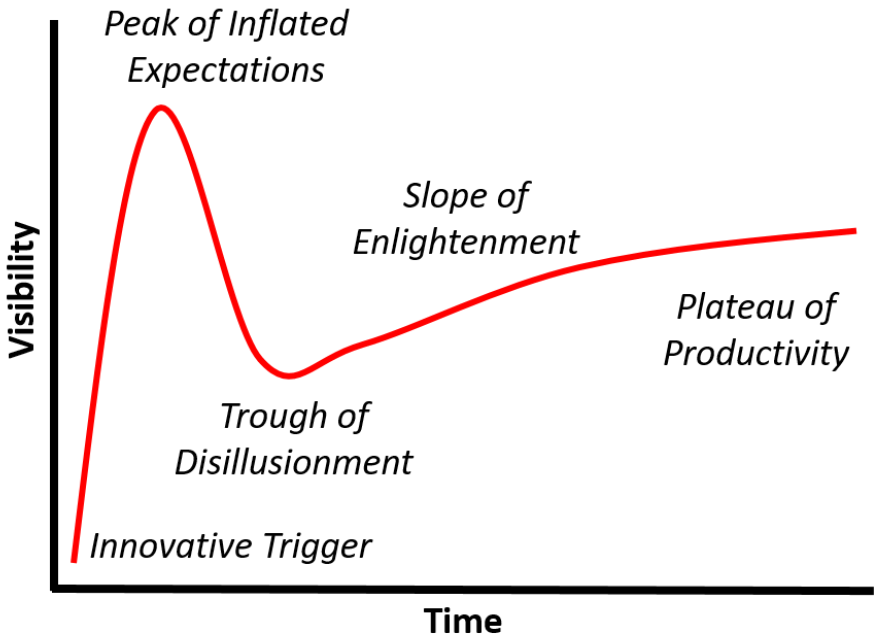


Figure 6.8 – Gartner's Hype Cycle, as described by Fenn and Rasinko (2008).

Chapter 7

Conclusions

7.1. Conclusions

This study presents an additional account of testing new methodological approaches on materials from the Olduvai Gorge. This iconic part of East African cultural heritage is still yielding a very high number of fossil remains, all from sites of increasing importance from multiple perspectives in the study of Human Evolution. FLK-West is one of the oldest known archaeological site presenting both lithic and faunal remains associated with the Acheulean technocomplex. Dating back to *circa* 1.7 Ma, this accumulation of lithic and osteological materials could be an important site that has the potential of revealing numerous features about these first Acheulean populations. The data that can be obtained from such an accumulation could shed light on the cognitive capabilities of these early hominins, while also providing valuable information for the reconstruction of the landscape they lived in. The formation processes responsible for the formation of FLK-West, however, are complex and have proven difficult to study in detail. This has been observed to be especially relevant regarding the faunal remains found within this accumulation. As has been revealed in the numerous preliminary studies regarding this site, the inflated equifinality observed product of both biostratinomic and diagenetic processes has meant that analysts require a more objective and new methodological toolkit in order to confront this site's interpretation.

This study set out on a search for alternatively new methods that could confront the archaeological and taphonomic record of sites such as this one. The initial objectives were adapted and moulded to the particular issue at hand, thus searching for a means of objectively and quantitatively classifying the different traces of FLK-West that had been distorted by other taphonomic processes. Since this initial objective was proposed, investigations performed over the last two years have evolved and changed to the different problems found throughout this body of research.

Conclusion 1

The first breakthrough can be found in the use of 3D digital microcopy for processing superficial marks such as trampling marks. The use of the HIROX KH-8700 3D Digital Microscope has enabled the efficient processing of taphonomic traces, especially those present at the archaeological site of FLK-West. We have also been able to refine the digital reconstruction protocol, aiding in a more efficient use of this piece of equipment. These results also gave fruit to the discovery of two different types of trampling marks observed within these experimental samples. Both traces have since been named *Scratches* and *Graze*s. Since the discovery of these variations in trampling marks, the results have further been able to refine our understanding of diagenetic processes, revealing the state of the bone and the consequent hardness of cortical surfaces to be a significant conditioning factor in trampling mark formation and morphology. A higher frequency of scratches is therefore an indication that trampling was produced while the bone was still fresh, while grazes are an indication that the bone was dry when the trampling phenomenon took place.

Conclusion 2

A successful comparison of trampling marks and cut marks was able to confirm previous known hypotheses regarding their differentiation, once the morphology of these different traces had been characterised. This research was able to reveal that the depth, curvature, width and length of these traces are important components in linear trace morphology.

Conclusion 3

The introduction of Artificial Intelligence algorithms has presented a significant improvement in our ability to process complicated data sets. Inspired by the recent introduction of Machine Learning in taphonomy, an exhaustive review of literature from all fields that employ statistics, ranging from archaeology to cancer research and pharmaceutical studies, was able to show how these algorithms work, and how they can be used to our benefit. With this, the introduction of Machine and Deep Learning for taphonomic Geometric Morphometric data was conceived, reaching 100% classification on training, validation and test results. When used to process archaeological traces, the different models predicted the class of each trace with an average of over 95% confidence. Furthermore, these results have been able to highlight how traditional discriminant functions such as LDA are not always powerful models when fit on some types of data. Carnivore tooth pits, for example, do not meet any of the mathematical requirements that LDA needs to differentiate between groups. Therefore, different classification methods are required in order to efficiently process certain databases.

Hybrid Artificial Intelligence and Geometric Morphometric approaches, therefore, could present a powerful paradigm shift that changes our understanding and ability to process the taphonomic register. While limitations do exist, future experimentation and developments of the present studies could be able to overcome the different issues that may arise from other sites. In as such, future research could provide a means of reaching higher degrees of resolution and more robust models for the differentiation of taphonomic traces, depending on the questions proposed by the analyst.

In general, this study has proposed a number of different means in which taphonomists can confront the fossil register. This can be considered an important basis on which to build upon, providing a methodological background that can be used to develop future research and investigations. While the methods developed were proposed and develop to respond to specific questions presented by the site of FLK-West, these methods can be extrapolated to almost any site, provided that the experimental sample being used for reference and comparison are truly analogous. With this, the results presented within this Master's thesis propose a new means to confronting taphonomic equifinality, potentially revealing highly important information for Human Evolution that had not been thought possible over a decade ago.

7.2. Prospective and Future Research

Firstly, the available techniques for processing microscopic data is still a key field of research. While this study has seen the potential of digital microscopes, in other parts of the world researchers are making valuable progress using Microtopography, Confocal and even 3D Scanning Electron Microscopy. In this study Microphotogrammetry, Structured Light Surface Scanning and Microscopy have been studied and compared statistically. Nevertheless, a combination of all of these approaches, in depth comparisons and statistical testing are needed before different methodological approaches can be made available on a wider scale. The availability of certain pieces of equipment is a significant problem in some centres of investigation, however this doesn't mean that there are no alternative approaches. The largest requirements are, therefore, an in depth comparison of the compatibility of these different pieces of equipment. Through this, regardless of whether the analyst has a 3D

digital microscope or a confocal profilometer, different teams need to be aware of the different possibilities each technique may have to offer and how they can be combined.

While this first reflection refers specifically to microscopy, the same point can be extrapolated to statistical approaches in archaeological data mining and data science as well. The results presented within this Master's thesis have highlighted how traditional discriminant functions such as LDA are not always powerful models when fit on some types of data. Carnivore tooth pits, for example, do not meet any of the mathematical requirements that LDA needs to differentiate between groups. These thus prove to be very poor performing classification models for archaeological data. With this proposal it is advisable that analysts try and evaluate multiple techniques before studying their final data, thus being open to new statistical methods that could prove much more reliable when extracting results.

Here it is advised that archaeologists ensure they are aware of the limitations of some statistical tests. It is also advisable that analysts update their methods in order to keep up with other scientific fields that are constantly evolving. This does not mean archaeologists should become well trained mathematicians, but does mean that before applying a statistical test, the user must consult the documentation available to ensure that there are no parameters that could be conditioning or providing erroneous results.

The next proposal builds off from this point, by advising analysts to look into other fields of research much more often and to keep an open mind towards the advances that fields such as medicine and engineering may have to offer. For legal as well as ethical reasons, most of these fields require incredibly strict statistical protocols and means of collecting data. For most doctors, the statistics being performed could be a matter of life or death. With this comes a much greater responsibility which needs more developed and advanced ways of looking at problem solving. These fields thus resort to searching for very powerful algorithms to confront such questions. In this study Random Forests and Machine Learning decision trees failed to provide useful results for the processing of morphological data, however this is not to say they may not prove valuable in lithic research, for example.

As an indication to the future of this line of research, as has been seen in FLK-West, while new methods are now available to confront the different issues present, this is not to say that the problem of equifinality has been completely solved. Future research into overlapping taphonomic traces is required, especially considering the incompleteness of some traces. The issue of missing landmarks, distorted traces and osteological surfaces that are hard to process require further research to truly overcome multiple issues that are still conditioning our abilities to study the taphonomic register in detail. This requires more experimental samples under different conditions, such as erosion by fluvial agents. Additionally, it would be useful to consider how mechanical cleaning processes and other restauration techniques may be distorting our data. Finally, more research is needed to see how the loss of landmarks may affect our interpretations and ability to study certain sites.

With this, the following years of research will be orientated around evaluating the current limitations of our methods and searching for ways to overcome them. Not only this, but the size of experimental samples will also need to increase, including more carnivore agents for African studies, such as different species of hyenas, leopards and African canids, and European studies, including foxes, bears and larger wolf samples. New experimentation will also be directed towards obtaining true analogy between both experimental and archaeological samples. This will include experimentation with fluvial abrasion, as well as overlapping traces that may alter our perception of the taphonomic register.

With this, the final note to make is to point out that the current study of FLK-West is only preliminary for the moment. Until the site

is expanded and fully excavated, no final conclusions can be withdrawn as to the taphonomic interpretation of the entire site. Future excavation campaigns in the years to come will hopefully provide more materials that can be studied and compared. With this, the hope is to provide a more detailed insight into the nature of hominin populations 1.7 million years ago. Nevertheless, this Master's Thesis still supports that the presented methods could be an important protagonist in future studies of the archaeological record; especially that associated with the earliest known Acheulean populations of Eastern Africa.

Chapter 8

Bibliography

Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; Kudlur, M.; Levenberg, J.; Monga, R.; Moore, S.; Murray, D.G.; Steiner, B.; Tucker, P.; Vasudevan, V.; Warden, P.; Wicke, M.; Yu, Y.; Zheng, X. (2016) TensorFlow: a system for large-scale machine learning. In: *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 2016), Savannah, GA, USA.* 16: 265–283. https://www.tensorflow.org

Abramowitz, M.; Davidson, M.W. (1999) Introduction to Microscopy, *Molecular Expressions*. [Online] Available from: http://micro.magnet.fsu.edu/primer/anatomy/introduction.html [Accessed: 26/02/2019]

Abrams, G; Bello, S.M.; Modica, K.D.; Pirson, S.; Bonjean, D. (2014) When Neanderthals used Cave Bear (*Ursus Spelaeus*) Remains: Bone Retouchers from Unit 5 of Scladina Cave (Belgium), *Quaternary International*. 326-327:274-287

Ackermann, R.R. (2002) Common Patterns of Facial Ontogeny in the Hominid Lineage, *The Anatomical Record*. 269:142-147.

Adams, D.C. (1999) Methods for Shape Analysis of Landmark Data from Articulated Structures, *Evolutionary Ecology Research*. 1:959-970

Adams, D.C.; Otárolla-Castillo, E. (2013) Geomorph: an R Package for the Collection and Analysis of Geometric Morphometric Shape Data, *Methods in Ecology and Evolution*. 4:393-399. DOI: 10.1111/2041-210X.12035

Adams, D.C.; Collyer, M.L.; Kaliontzopoulou, A.; Sherratt, E. (2017) Geomorph: Software for Geometric Morphometric Analysis. R Package Version 3.0.5. http://cran.r-project.org/package=geomorph

Agam, A.; Barkai, R. (2018) Elephant and Mammoth Hunting during the Palaeolithic: a Review of the Relevant Archaeological, Ethnographic and Ethno-Historical Records, *Quaternary*. 1(3):1-28 DOI: 10.3390/quat1010003

Aiello, L.C.; Wheeler, P. (1995) The Expensive Tissue Hypothesis, *Current Anthropology*. 36:199-221

Albrecht, G.H. (1992) Assessing the Affinities of Fossils using Canonical Variates and Generalized Distances, *Journal of Human Evolution*. 7(4):49-69

Albrecht, G.H. (1992) Assessing the Affinities of Fossils using Canonical Variates and Generalized Distances, *Human Evolution*. 7(4):49-69

Almécija, S.; Moyà-Solà, S.; Alba D.M. (2010) Early Origin for Human-Like Precision Grasping: A Comparative Study of Pollical Distal Phalanges in Fossil Hominins, *PLoS ONE*. DOI: 10.1371/journal.pone.0011727

Anaconda Software Distribution (2016) Anaconda Computer Software. https://anaconda.com

Andrés, M; Gidna, AO; Yravedra, J; Domínguez-Rodrigo, M (2012) A Study of Dimensional Differences of Tooth Marks (Pits and Scores) on Bones Modified by Small and Large Carnivores, *Journal of Archaeological and Anthropological Sciences*. 4(3):209-219

Andrews, P.; Cook, J. (1985) Natural Modifications to Bones in a Temperate Setting, *Man*. 20(4):675-691

Andrews P, Fernandez-Jalvo Y (1997) Surface modifications of the Sima de los Huesos fossil humans. *Journal of Human Evolution* 33:191–217.

Aramendi, J.; Maté-González, M.A.; Yravedra, J.; Cruz Ortega, M.; Arriaza, M.C.; González-Aguilera, D.; Baquedano, E.; Domínguez-Rodrigo M. (2017a) Discerning carnivore agency through the three-dimensional study of tooth pits: revisiting crocodile feeding leistoc at FLK-Zinj and FLK NN3 (Olduvai Gorge, Tanzania). *Palaeogeography, Palaeoclimatology, Palaeoecology*. 488:93-102 DOI: 10.1016/j.palaeo.2017.05.021

Aramendi, J.; Uribelarrea, D.; Arriaza, M.C.; Arráiz, H.; Barboni, D.; Yravedra, J.; Ortega, M.C.; Gidna, A.; Mabulla, A.; Baquedano, E.; Domínguez-Rodrigo (2017b) The Paleoecology and Taphonomy of AMK (Bed I, Olduvai Gorge) and its Contributions to the Understanding of the "Zinj" Paleolandscape, *Palaeogeography, Palaeoclimatology, Palaeoecology*. 488:35-49 DOI: 10.1016/j.palaeo.2017.02.036

Araujo, A.G.M.; Marcelino, J.C. (2003) The Role of Armadillos in the Movement of Archaeological Materials: an Experimental Approach, *Geoarchaeology: an International Journal*. 18(4):433-460

Arbour, J.H.; Brown, C.M. (2014) Incomplete Specimens in Geometric Morphometric Analysis, *Methods in Ecology and Evolution*. 5:16-26. DOI: 10.1111/2041-210X.12128

Archer, W.; Braun, D.R. (2013) Investigating the Signature of Aquatic Resource Use within Pleistocene Hominin Dietary Adaptations, *PLoS One*. 8(8):e6899 DOI: 10.1371/journal.pone.0069899

Arilla, M.; Rosell, J.; Blasco, R.; Domínguez-Rodrigo, M.; Pickering, T.R. (2014) The "Bear" Essentials: Actualistic Research on *Ursus arctos arctos* in the Spanish Pyrenees and its Implications for Paleontology and Archaeology, *PLoS One*. 9(7):e102457 DOI: 10.1371/journal.pone.0102457

Arnqvist, G.; Mártensson, T. (1998) Measurement Error in Geometric Morphometrics: Empirical Strategies to Assess and Reduce its Impact on Measures of Shape, *Acta Zoologica Academiae Scientiarum Hungaricae*. 44(1):73-96

Arráiz, H.; Barboni, D.; Uribellarea, D.; Mabulla, A.; Baquedano, E.; Domínguez-Rodrigo, M. (2017) Paleovegetation Changes Accompanying the Evolution of a Riverine System at the BK Paleoanthropological Site (Upper Bed II, Olduvai Gorge, Tanzania), *Palaeogeography, Palaeoclimatology, Palaeoecology.* 488:84-92. DOI: 10.1016/j.palaeo.2017.05.010

Arriaza, M.C.; Domínguez-Rodrigo, M.; Yravedra, J.; Baquedano, E. (2016) Lions as Bone Accumulators? Paleontological and Ecological Implications of a Modern Bone Assemblage from Olduvai Gorge. *PLoS ONE*. 11:e0153797. DOI: 10.1371/journal.pone.0153797

Arriaza, M.C.; Domínguez-Rodrigo, M. (2016) When Felids and Hominins ruled at Olduvai Gorge: a Machine Learning Analysis of Skeletal Profiles of the Non-Anthropogenic Bed I Sites, *Quaternary Science Reviews*. 139:43-52

Arriaza, M.C.; Yravedra, J.; Domínguez-Rodrigo, M.; Maté-González, M.Á.; Vargas, E.G.; Palomeque-González, J.P.; Aramendi, J.; González-Aguilera, D.; Baquedano, E. (2017) On Applications of Micro-Photogrammetry and Geometric Morphometrics to Studies of Tooth Mark Morphology: The Modern Olduvai Carnivore Site (Tanzania). *Palaeogeography, Palaeoclimatology, Palaeoecology*. DOI: 10.1016/j.palaeo.2017.01.036

Arribas, A.; Palmqvist, P. (1999) On the Ecological Connection between Sabre-Tooths and Hominids: Faunal Dispersal Events in the Lower Pleistocene and a Review of the Evidence for the First Human Arrival in Europe, *Journal of Archaeological Science.* 26:571-585

Artheya, S. (2012) The Frontal Bone in the Genus *Homo*: a Survey of Functional and Phylogenetic Sources of Variation, *Journal of Anthropological Sciences.* 90:59-80

Ashley, G.M. (2007) Orbital Rythms, Monsoons, and Playa Lake Response, Olduvai Basin, Equatorial Africa (1.85-1.75), *Geology*. 35:1091-1094.

Ashley, G.M.; Barboni, D.; Domínguez-Rodrigo, M.; Bunn, H.T.; Mabulla, A.Z.P.; Diez-Martin, F.; Barba, R.; Baquedano, E. (2010) A Spring and Wooded Habitat at FLK Zinj and their Relevance to Origins of Human Behaviour, *Quaternary Research*. 74:304-314.

Atici, L. (2006) Middle-Range Theory in Paleolithic Archaeology: The Past and the Present, *Journal of Taphonomy*. 4(1):29-45.

Baab, K.L.; McNulty, K.P.; Rohlf, F.J. (2012) The Shape of Human Evolution: A Geometric Morphometrics Perspective, *Evolutionary Anthropology*. 21:151-165

Baddeley, A.; Gegori, P.; Mahiques, J.M.; Stoica, R.; Stoyan, D. (2001) *Case Studies in Spatial Point Modeling*. The Netherlands: Springer.

Baquedano, E; Domínguez-Rodrigo, M; Musiba, C (2012) An Experimental Study of Large Mammal Bone Modification by Crododiles and its Bearing on the Interpretation of Crocodile Predation at FLK Zinj and FLK NN3, *Journal of Archaeological Science*. 39(6):1728-1737. DOI: 10.1016/j.jas.2012.01.010

Barboni, D.; Ashley, G.M.; Domínguez-Rodrigo, M.; Bunn, H.T.; Mabulla, A.Z.P.; Baquedano, E. (2010) Phytoliths Infer Locally Dense and Heterogeneous Paleovegetation at FLK North and Surrounding Localities during Upper Bed I Time, Olduvai Gorge, Tanzania, *Quaternary Research*. 74(3):344-354

Bartosiewicz, L. (2009) Skin and Bones: Taphonomy of a Medieval Tannery in Hungary. *Journal of Taphonomy* 7(2):91-107.

Bartram, L.E.; Marean, C.W. (1999) Explaining the "Klasies Pattern": Kua Ethnoarchaeology, the Die Kelders Middle Stone Age Archaeofauna, Long Bone Fragmentation and Carnivore Ravaging, *Journal of Archaeological Science*. 26:9-29

Bastir, M.; Rosas, A. (2004a) Comparative Ontogeny in Humans and Chimpanzees: Similarities, Differences and Paradoxes in Postnatal Growth and Development of the Skull, *Annals of Anatomy*. 186:503-509

Bastir, M; Rosas, A. (2004b) Facial Heights: Evolutionary Relevance of Postnatal Ontogeny for Facial Orientation and Skull Morphology in Humans and Chimpanzees, *Journal of Human Evolution*. 47:359-381

Bastir, M.; Rosas, A.; O'Higgins, P. (2006) Craniofacial Levels and the Morphological Maturation of the Human Skull, *Journal of Anatomy*. 209:637-654. DOI: 10.1111/j.1469-7580.2006.00644.x

Bastir, M.; O'Higgins, P.; Rosas, A. (2007) Facial Ontogeny in Neanderthals and Modern Humans, *Proceedings of the Royal Society B*. 274:1125-1132 DOI: 10.1098/rsqb.2006.0448

Basu, S.; Mukhopadhyay, S.; Karki, M.; DiBiano, R.; Ganguly, S.; Nemani, R.; Gayaka, S. (2018) Deep Neural Networks for Texture Classification – A Theoretical Analysis, *Neural Networks*. 97:173-182

Beauval, C.; Morin, E. (2010) Les Repaires d'Hyènes du Lussacois (Lussac-les-Châteaux, Vienne, France): Apport des Sites des Plumettes et des Rochers-de-Villeneuve. In: J. Buisson-Catil and J. Primaul (Eds.) *Préhistoire entre Vienne et Charente*. Chauvigny: Association des Publications Chauvinoises.

Beerenwinkel, N.; Däumer, M.; Oette, M.; Korn, K.; Hoffmann, D.; Kaiser, R.; Lengauer, T.; Selbig, J.; Walter, H. (2003) Geno2pheno: Estimating Phenotypic Drug Resistance from HIV-1 genotypes, *Nucleic Acids Res.* 31:3850-3855

Beerenwinkel, N.; Schmidt, B.; Walter, H.; Kaiser, R.; Lengauer, T.; Hoffmann, D.; Korn, K.; Selbig, J. (2002) Diversity and Complexity of HIV-1 Drug Resistance: Bioinformatics Approach to Predicting Phenotype from Genotype, *PNAS*. 99:8271-8276

Behrensmeyer, A.K. (1978) Taphonomic and Ecologic Information from Bone Weathering. *Paleobiology*. 4(2):150-162.

Behrensmeyer, A.K.; Gordon, K.D.; Yanagi, G.T. (1986) Trampling as a Cause for Bone Surface Damage and Pseudo-Cutmarks, *Nature*. 319:768-771.

Behrensmeyer, A.K.; Kidwell, S.M.; Gastaldo, R.A. (2000) Taphonomy and Paleobiology, *Paleobiology*. 26(4):103-147.

Bellazzi, R.; Zupan, B. (2008) Predictive Data Mining in Clinical Medicine: Current Issues and Guidelines, *International Journal of Medical Information.* 77(2):81-97

Bello, S.M.; Soligo, C. (2008) A new method for the quantitative analysis of cutmark micromorphology. *Journal of Archaeological Science* 35(6):1542–1552.

Bello, S.; Parfitt, S.; Stringer, C. (2009) Quantitative micromorphological analyses of cut marks produced by ancient and modern handaxes, *Journal of Archaeological Science* 36(9):1869–1880.

Bello, S. (2011) New Results from the Examination of Cut-Marks using Three-Dimensional Imaging. In: Ashoton, N.M., Lewis, S.G. and Stringer, C.B. (eds.) *The Ancient Human Occupation of Britain*. Amsterdam: Elsevier. 249-262

Bello, S; Groote, I.; Delbarre, G. (2013) Application of 3-dimensional microscopy and micro-CT scanning to the analysis of Magdalenian portable art on bone antler. *Journal of Archaeological Science.* 40(5):2464–2476

Bello, SM; Wallduck, R; Dimitrijevic, V; Zivaljevic, I; Stringer, CB (2016) Cannibalism Versus Funerary Defleshing and Disarticulation after a Period of Decay: Comparisons of Bone Modifications from Four Prehistoric Sites, *American Journal of Physical Anthropology*. DOI: 10.1002/ajpa.23079

Bello, S.; Galway-Witham, J. (2019) Bone Taphonomy Inside and Out: Application of 3-Dimensional Microscopy, Scanning Electron Microscopy and Micro-Computed Tomography to the Study of Humanly Modified Faunal Assemblages, *Quaternary International*. DOI: 10.1016/j.quaint.2019.02.035

Benazzi, S.; Stansfield, E.; Kullmer, O.; Fiorenza, L.; Gruppioni, G. (2009) Geometric Morphometric Methods for Bone Reconstruction: The Mandibular Condylar Process of Pico della Mirandola, *The Anatomical Record*. 292:1088-1097.

Bertalanffy, L. (1949) *Problems of Life: an Evaluation of Modern Biological and Scientific Thought*. New York: Harper.

Bertalanffy, L. (1698) Types of Finality , *General System Theory*. New York: George Braziller. 77-80.

Beyene, Y.; Katoh, S.; WoldeGabriel, G.; Hart, W.K.; Uto, K.; Sudo, M.; Kondo, M.; Hyodo, M.; Renne, P.R.; Suwa, G.; Asfaw, B. (2012) The Characteristics and Chronology of the Earliest Acheulean at Konso, Ethiopia, *PNAS*. 110(5):1584-1591

Bilingsley, H. (1570) *Euvlid's Elements*. London: John Dave.

Binford, L.R. (1962) Archaeology as Anthropology, *American Antiquity*. 28(2):217-225

Binford, L.R. (1967) Smudge Pits and Hide Smoking: The Use of Analogy in Archaeological Reasoning, American Antiquity. 32(1):1-12

Binford, L.R. (1968) Archaeological Perspectives. In: S.R. Binford and L.R. Binford (Eds.) *New Perspectives in Archaeology*. Aldine, New York. 5-32

Binford, L.R. (1978) *Nunamiut Ethnoarchaeology*. New York: Academic Press Inc.

Binford, L.R. (1981) *Bones: Ancient Men and Modern Myths*. New York: Academic Press Inc.

Binford, L.R. (1983) Reply to "More on the Mousterian Flaked Bone from Cueva Morín" by L. Freeman, *Current Anthropology*. 24:372-377

Binford, L.R. (1984a) *Faunal Remains from Klasies River Mouth.* New York: Academic Press Inc.

Binford, L.R. (1984b) Butchering, Sharing and the Archaeological Record, *Journal of Anthropological Archaeology*. 3:235-257

Binford, L.R. (1985) Human Ancestors: Changing Views of their Behavior, *Journal of Anthropological Archaeology*. 4:292-327

Binford, L.R. (1987) Were there Elephant Hunters at Torralba? In: M.H. Nitecki, D.W. Nitecki (Eds.) *The Evolution of Human Hunting*. Boston: Springer. 47-104

Binford, L.R.; Stone, N.M. (1987a) On Inference from the Zhoukoudian Fauna, *Current Anthropology*. 28(3):358-362

Binford, L.R.; Stone, N.M. (1987b) On Zhoukoudian, *Current Anthropology*. 28(1):102-105

Binford, L.R. (1988a) Fact and Fiction about the Zinjanthropus Floor, Data, Arguments and Interpretations, *Current Anthropology*. 29(1):123-135

Binford, L.R. (1988b) Etude Taphonomique des Restes Fauniques de la Grotte Vaufrey. In: J.P. Rigaud (Ed.) *La Grotte Vaufrey à Cénac et Sain Julien (Dordogne): Paléoenvironments, Chronologie et Activités Humaines*. Paris: Mémoires de la Société Préhistorique Française. 19:535-564.

Binford, L.R. (1988c) The Hunting Hypothesis: Archaeological Methods and the Past, *Yearbook of Physical Anthropology*. 30:1-9

Binford, L.R. (1989) Letter to H.T. Bunn, *Debating Archaeology.* Abingdon: Routledge. 291-293

Bishop, W.; Clark, J.D. (1967) *Background to Evolution in Africa*. Chicago: University of Chicago Press.

Bishop, C. (1995) *Neural Networks for Pattern Recognition*. New York: Oxford University Press

Bishop, C. (2006) *Pattern Recognition and Machine Learning*. Singapore: Springer.

Bisoi, R.; Dash, P.K.; Parida, A.K. (2019) Hybrid Variational Mode Decomposition and Evolutionary Robust Kernel Extreme Learning Machine for Stock Price and Movement Prediction on Daily Basis, *Applied Soft Computing Journal*. 74:652-678

Blasco, R.; Rosell, J.; Fernández Peris, J.; Cáceres, I.; Vergès, J.M. (2008) A New Element of Trampling: an Experimental Application on the Level XII Faunal Record of Bolomor Cave (Valencia, Spain), *Journal of Archaeological Science*. 35:1605-1618

Blasco, R.; Rosell, J. (2009) Who was the first? An Experimental Application of Carnivore and Hominid Overlapping Marks at the Pleistocene Archaeological Sites, *Palevol*. 8:579-592

Blasco, R.; Rosell, J.; Arsuaga, J.L.; Bermúdez de Castro, J.M.; Carbonell, E. (2010) The Hunted Hunter: the Capture of a Lion (*Panthera leo fossilis*) at the Gran Dolina Site, Sierra de Atapuerca, Spain, *Journal of Archaeological Science*. 37:2051-2060

Blasco, R.; Domínguez-Rodrigo, M.; Arilla, M.; Camarós, E.; Rosell, J. (2014) Breaking Bones to Obtain Marrow: a Comparative Study Between Percussion by Batting Bone on an Anvil and Hammerstone Percussion, *Archaeometry*. DOI: 10.1111/arcm.12084

Blasco, R.; Rosell, J.; Smith, K.T.; Maul, L.C.; Sañudo, P.; Barkai, R.; Gopher, A. (2016) Tortoises as a Dietary Supplement: a View from the Middle Pleistocene Site of Qesem Cave, Israel, *Quaternary International Science Reviews*. 133:165-182

Blue Marble Geographics (2019) Global Mapper Software. v.15. https://www.bluemarblegeo.com/products/global-mapper.php

Blumenschine, R. (1986) *Early Hominid Scavenging Opportunities. Implications of Carcass Availability in the Serengeti and Ngorongoro Ecosystems*. Oxford: Archaeopress

Blumenschine, R. (1988) An Experimental Model of the Timing of Hominid and Carnivore Influence on Archaeological Bone Assemblages, *Journal of Archaeological Science.* 15, 483-502

Blumenschine, R.J.; Selvaggio, M.M. (1991) On the marks of marrow bone processing by hammerstones and hyaenas: their anatomical patterning and archaeological implications. In J.D. Clarke (Ed.) Cultural Beginnings: Approaches to Understanding Early Hominid Life Ways in the African Savanna. Switzerland: R. Habelt. 17-32

Blumenschine, R.; Marean, C.W. (1993) Long Bone Marrow Yields of Some African Ungulates, *Journal of Archaeological Science*. 20:555-587

Blumenschine, R. (1995) Percussion Marks, Tooth Marks and Experimental Determinations of the Timing of Hominid and Carnivore Access to Long Bones at FLK Zinjanthropus, Olduvai Gorge, Tanzania. *Journal of Human Evolution*. 29(1):21-51

Blumenschine, R.; Marean, C.W.; Capaldo, S. (1996) Blind Tests of Inter-Analyst Correspondence and Accuracy in the Identification of Cut Marks, Percussion Marks, and Carnivore Tooth Marks on Bone Surfaces, *Journal of Archaeological Science*. 23:493-507

Blumenschine, R: Prassack, KA; Kreger, D; Pante, MC (2007) Carnivore Tooth-Marks, Microbial Bioerosion, and the Invalidation of Domínguez-Rodrigo and Barba's (2006) Test of Oldowan Hominin Scavenging Behavior, *Journal of Human Evolution*. 53, 420-426.

Blumenschine, R.J.; Masao, F.T.; Tactikos, J.C.; Ebert, J.I. (2008) Effects of distance from stone source on landscape-scale variation in Oldowan artifact assemblages in the paleo-Olduvai Basin, Tanzania. *Journal of Archaeological Science*. 35:76–86

Blumenschine, R.J.; Stanistreet, I.G.; Njau, J.K.; Bamford, M.K.; Masao, F.T.; Albert, R.M.; Stollhofen, H.; Andrews, P.; Prassack, K.A.; McHenry, L.J.; Fernández-Jalvo, Y.; Camilli, E.L.; Ebert, J.I. (2012) Environments and Hominin Activities Across the FLK Peninsula During *Zinjanthropus* times (1.84 Ma), Olduvai Gorge, Tanzania, *Journal of Human Evolution*. 63:364-383.

Boaz, N.T.; Ciochon, R.L.; Xu, Q.; Liu, J. (2004) Mapping and Taphonomic Analysis of the *Homo erectus* at Locality 1 Zhoukoudian, China, *Journal of Human Evolution*. 46:519-549

Böhmer, K.; Rinderle-Ma, S. (2016) Multi-Perspective Anomaly Detection in Business Process Execution Events. *On the Move to Meaningful Internet Systems*. The Netherlands: Springer 80-98.

Bonhomme, V.; Picq, S.; Gaucherel, C.; Claude, J. (2014) Momocs: Outline Analysis using R, *Journal of Statistical Software,* 56(13):1-24

Bonifay, MF (1996) Les Carnivores. In Lavocat, R (Ed.) *Faunes et Flores Préhistoriques de l'Europe Occidentale.* Paris: Centre National de la Recherche Scientifique 3:57-58

Bonney H (2014) An investigation of the use of discriminant analysis for the classification of blade edge type from cut marks made by metal and bamboo blades. *American Journal of Physical Anthropology.* 154:575–584

Bookstein, F.L. (1989) Principal Warps: Thin Plate Spline and the Decomposition of Deformations. *Transactions on Pattern Analysis and Machine Intelligence*. 11(6):567-585

Bookstein, F.L. (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. New York: Cambridge University Press.

Bookstein, F.L. (1996) Landmark Methods for Forms without Landmarks: Morphometrics of Group Differences in Outline Shape, *Medical Image Analysis*. 1:225-243.

Bookstein, F.L. (1998) A Hundred Years of Morphometrics, *Acta Zoologica Academiae Scientarium Hungaricae*. 44(1):7-59

Bookstein, F.L.; Shafer, K.; Prossinger, H.; Seidler, H.; Fieder, M.; Stringer, C.; Weber, G.W.; Arsuage, J.L.; Slice, D.E.; Rohlf, F.J.; Recheis W.; Mariam, A.J.; Marcus, L.F. (1999)

Comparing Frontal Cranial Profiles in Archaic and Modern *Homo* by Morphometric Analysis, *The Anatomical Record*. 257:217-224

Bordes, F. (1988) *Typologie du Paléolithique Ancien et Moyen*. Paris: CNRS

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Boschin, F.; Crezzini, J. (2011) Morphometrical Analysis on Cut Marks using a 3D Digital Microscope, *International Journal of Osteoarchaeology.* DOI: 10.1002/oa.1272

Bradbury, S.; Bracegirdle, B. (1998) *Introduction to Light Microscopy*. Milton Park: BIOS Science.

Brande, S.; Saragusti, I (1999) Arti-Facts Graphic Visualization of Handaxes and Other Artifacts, *Near Eastern Archaeology*. 62(4):242-245

Brain, C.K. (1967) Bone Weathering and the Problem of Pseudo-Tools, *South African Journal of Science*. 63:97-99.

Brain, C.K. (1969) *Faunal remains from the Wilton Large Rock Shelter. In Deacon, J., Re-excavation and Description of the Wilton Type-site, Albany District, Eastern Cape*. Masters Thesis. Cape Town: University of Cape Town.

Brain, C.K. (1974) Some Suggested Procedures in the Analysis of Bone Accumulations from Southern African Quaternary Sites. *Annals of the Transvaal Museum*. 29: 1-5

Brain, C.K. (1981) *The Hunters or the Hunted?: An Introduction to African Cave Taphonomy*. Chicago: The University of Chicago Press.

Braun, D.R.; Pobiner, B.L.; Thompson, J.C. (2008) An Experimental Investiagation of Cut Mark Production and Stone Tool Attrition, *Journal of Archaeological Science*. DOI: 10.1016/j.jas.2007.08.015

Braun, D.R.; Pante, M.; Archer, W. (2016) Cut Marks on Bone Surfaces: Influences on Variation in the Form of Traces of Ancient Behaviour, *Interfoace Focus*. DOI: 10.1098/rsfs.2016.0006

Breiman, L.; Friedman, J.H.; Stone, C.J.; Olshen, R.A. (1984) *Classification and Regression Trees*. London: Chapman and Hall.

Breiman, L. (2001) Random Forests, *Machine Learning*. 45:5-32

Breiman, L.; Cutler, A. (2018) Package 'randomForest': Breiman and Cutler's Random Forests for Classification and Regression. R Package Version 4.6-14 https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Bromage, T.G.; McMahon, J.M.; Thackeray, J.F.; Killmer, O.; Hogg, R.; Rosenberger, A.L.; Schrenk, F.; Enlow, D.H. (2008) Craniofacial Architectural Constraints and their Importance for Reconstructing the Early *Homo* skull KNM-ER 1470, *Journal of Clinical Pedratic Dentistry*. 33(1):43-54

Brownlee, J. (2016a) *Machine Learning Mastery with R*. Melbourne: Machine Learning Mastery.

Brownlee, J. (2016b) *Deep Learning with Python*. Melbourne: Machine Learning Mastery.

Brownlee, J. (2019a) *Better Deep Learning: Train Faster, Reduce Overfitting and Make Better Predictions*. Melbourne: Machine Learning Mastery

Brownlee, J. (2019b) *Deep Learning for Computer Vision*. Melbourne: Machine Learning Mastery

Brugal, J.P. (1995) Middle Palaeolithic Subsistence on Large Bovids: La Borde Coudolous 1 (Lot, France): Problems and Methods. In: S. Gaudzinski, S. and E. Turner (Eds.) *The Roles of Early Humans in the Accumulations of European Lower and Middle Palaeolithic Bone Assemblages*. Neuwied: Foschungsbereich Altseinzeit des Römisch-Garemanischen Zentralmuseums Mainz. 30-31

Burgal, J.P.; Fosse, P.; Guagelli, J.L. (1997) Comparative Study of Bone Assemblages made vt Recent Pleistocene Hyenids. In: L.A. Hannus, L. Rossum and R.P. Winham (Eds.) *Proceedings of the 1993 Bone Conference Hot Springs, South Dakota*. Sioux Falls: Archaeology Laboratory Augustana Collega. 158-187

Bruner, E.; Saracino, B.; Ricci, F.; Tafuri, M.; Passarello, P.; Manzi, G. (2004) Midsagittal Cranial Shape Variation in the Genus by Geometric Morphometrics, *Coll. Antropol.* 28(1):99-112

Bruner, E.; Athreya, S.; Cuétara, J.M.; Marks, T. (2013) Geometric Variation of the Frontal Squama in the Genus *Homo*: Frontal Bulging and the Origin of Modern Human Morphology, *American Journal of Physical Anthropology*. 150:313-323 DOI: 10.1002/ajpa.22202

Bui, H.M.; Lech, M.; Cheng, E.; Neville, K.; Burnett, I.S. (2016) Using Grayscale Images for Object Recognition with Convolutional-Recursive Neural Network, *International Conference on Communications and Electronics*. 6:321-325

Bunge, M. (1981) Analogy between Systems, *International Journal of General Systems.* 7:221-223. DOI: 10.1080/03081078108934823

Bunn, H.T.; Harris, J.W.K.; Isaac, G.; Kaufulu, Z.; Kroll, E.; Schick, K.; Toth, N.; Behrensmeyer, A. (1980) FxJj50: an Early Pleistocene Site in Northern Kenya, *World Archaeology*. 12(2):109-136

Bunn, H.T. (1981) Archaeological Evidence for Meat Eating by Plio-Pleistocene Hominids from Koobi Fora and Olduvai Gorge, *Nature*. 291(5816):574-577

Bunn, H.T. (1982a) *Meat Eating and Human Evolution: Studies on the Diet and Subsistence Patterns of Plio-Pleistocene Hominids in East Africa*. PhD Thesis. Berkley: University of California.

Bunn, H.T. (1982b) Animal Bones and Archaeological Inference, *Science*. 215:494-495

Bunn, H.T. (1983a) Evidence on the Diet and Subsistence Patterns of Plio-Pleistocene Hominids at Koobi Fora, Kenya, and Olduvai Gorge, Tanzania. In: J. Clutton-Bruck, C. Grigson (Eds.) *Animals and Archaeology: Hunters and their Prey.* BAR International Series 163:21-30

Bunn, H.T. (1983b) Comparative Analysis of Modern Bone Assemblages from a Sam Hunter-Gatherer Camp in a Kalahari Desert Botswana, and from Spotted Hyaena Den Near Nairobi, Kenya. In: J. Clutton-Brock and C. Grigson (Eds.) *Animals and Archaeology. Vol 1: Hunters and their Prey*. Oxford:BAR 163:143-148.

Bunn, H.T. (1986) Patterns of Skeletal Representation and Hominid Subsistence Activities at Olduvai Gorge, Tanzania and Koobi Fora, Kenya, *Journal of Human Evolution*. 15:673-690

Bunn, H.T.; Kroll, E.M. (1986) Systematic Butchery by Plio-Pleistocene Hominid at Olduvai Gorge, Tanzania, *Currenty Anthropology*. 27:431-452

Bunn, H.T. (1987) Patterns of Skeletal Representation and Hominid Subsistence Activities at Olduvai Gorge, Tanzania, and Koobi Fora, Kenya, *Journal of Human Evolution*. 15:673-690

Bunn, H.T.; Kroll, E.M. (1987) On Inference from the Zoukoudian Fauna, *Current Anthropology*. 28(3):199-202

Bunn, H.T.; Kroll, E.M. (1988) A Reply to Binford, *Current Anthropology*. 29:123-149

Bunn, H.T. (1989) Diagnosing Plio-Pleistocene Hominid Activity Bone Fracture Evidence. In: R. Bonnichsen and M. Sorg (Eds.) *Bone Modifications*. 299-316. Maine: Orono.

Bunn, H.T. (1993) Bone Assemblages at Base Camps: a Further Consideration of Carcass Transport and Bone Destruction by the Hadza. In J. Hudson (Ed.) *From Bones to*

Behaviour. Ethnoarchaeological and Experimental Contributions to the Interpretation of Faunal Remains. Illinois: University of Carbondale. 21:156-168

Bunn, H.T.; Edzo, J.A. (1993) Hunting and Scavenging by Pliocene Hominids: Nutritional Constraints, Archaeological Patterns, and Behavioural Implications, *Journal of Archaeological Science*. 20:365-398

Bunn, H.T. (1994) Early Pleistocene Hominid Foraging Strategies along the Ancestral Omo River at Koobi Fora, Kenya, *Journal of Human Evolution*. 27:247-266

Bunn, H.T. (1995) Reply to Tappen. *Current Anthropology*. 36:250-251

Bunn, H.T. (1996) Reply to Rose and Marshall, *Current Anthropology*. 37:321-323

Bunn, H.T. (2001) Hunting, Power-Scavenging and Butchering by Hadza Foragers by Plio-Pleistocene Homo. In: C. Stanford, H.T. Bunn (Eds.) *Meat Eating and Human Evolution*. Oxford: Oxfard University Press. 199-218

Bunn, H.T.; Pickering, T.R. (2010) Methodological Recommendations for Ungulate Mortality Analyses in Paleoanthropology, *Quaternary Research*. 74:388-394

Byeon, W.; Domínguez-Rodrigo, M.; Arampatzis, G.; Baquedano, E.; Yravedra, J.; Maté-González, M.A.; Koumoutsakos, P. (2019) Automated Identification and Deep Classification of Cut Marks on Bones and its Palaeoanthropological Implications, *Journal of Computational Science*. 32:36-43 DOI: 10.1016/j.jocs.2019.02.005

le Cabec, A.; Toussaint, M. (2017) Impacts of Curatorial and Research Practices on the Preservation of Fossil Hominid Remains, *Journal of Anthropological Sciences*. 95:1-28

Canty, A.; Ripley, B. (2017) Boot: Bootstrap R (S-Plus) Functions. R Package version 1.3-20.

Camarós, E.; Cueto, M.; Teira, L.; Tapia, J.; Cubas, M.; Blasco, R.; Rosell, J.; Rivals, F. (2013) Large Carnivores as Taphonomic Agents of Space Modification: an Experimental Approach with Archaeological Implications, *Journal of Archaeological Science*. 40:1361-1368

Capaldo, S.; Blumenschine, R.J. (1994) A Quantitative Diagnosis of Notches Made by Hammerstone Percussion and Carnivore Gnawing on Bovid Long Bones, *Society for American Archaeology*. 59(4):724-748.

Capaldo, S. (1995) Inferring Hominid and Carnivore Behaviour from Dual Patterned Archaeological Assemblages. PhD Thesis. New Brunswick: Rutgers University.

Capaldo, S. (1997) Experimental Determinations of Carcass Processing by Plio-Pleistocene Hominids and Carnivores at FLK 22 (Zinjanthropus) Olduvai Gorge, Tanzania. Journal of Human Evolution 33:555-597

Capaldo, S. (1998) Simulating the Formation of Dual-Patterned Archaeofaunal Assemblages with Experimental Control Simples, *Journal of Archaeological Science*. 35:311-330

Castelló, J.R. (2016) *Bovids of the World*. Oxford: Princeton University Press.

Carbonell, E.; Guilbaud, M.; Mora, R. (1983) Utilización de la Lógica Analitica para el Estudio de Tecnocomplejos a Cantos Tallados. *Cahier Noir* 1:3-64

Carbonell, E.; Mosquera, M.; Ollé, A.; Rodríguez, X.P.; Sala, R.; Vaquero, M.; Vergés, J.M. (1992) New Elements of the Logical Analytic System, *Cahier Noir*. 6:3-61

Cardillo, M. (2006) Explorando la Variación en las Morfologías Líticas a Partir de las Técnicas de Análisis de Contornos. El Caso de las Puntas de Proyectil del Holocenomedio-tardío de la Punta de Salta (San Antonio de los Cobres, Argentina), *Werken*. 7:77-88

Cardillo, M.; Charlin, J. (2009) Tendencias observadas en la variabilidad de los raspadores de norte y sur de Patagonia. Explorando las interrelaciones entre forma, tamaño e historia de vida. *Arqueometría Latinoamericana*. Buenos Aires: Congreso Argentino y lro Latinoamericano. 2(2):351–359.

Cardillo, M. (2009) *Variabilidad en la manufactura y diseño de artefactos en el área costera patagónica. Un enfoque integrador*. PhD Thesis. Buenos Aires: Universidad de Buenos Aires.

Cardillo, M. (2010) Some Applications of Geometric Morphometrics to Archaeology. In: A.M.T. Elewa (Ed.) *Morphometrics to Nonmorphometricians*. New York: Springer. 325-341

Carlson, D.S. (2005) Theories of Craniofacial Growth in the Postgenomic Era, *Orthodontics*. DOI: 10.1053/j.sodo.2005.07.002

Castiñeira, C.; Cardillo, M.; Chalin, J.; Fernicola, J.C.; Baeza, J. (2009) Análisis morfométrico de cabezales líticos "cola de pescado" de la Rep Oriental del Uruguay. *Arqueometría Latinoamericana*. Buenos Aires: Congreso Argentino y 1ro Latinoamericano. 2(1):360-366

CELSYS inc (2014) ClipStudio Pain, v. 1.8.4. https://www.clipstudio.net/en

Cela Conde, C.J.; Ayala, F.J. (2013) La Caune de l'Arago, *Evolución Humana: El Camino Hacia Nuestra Especie.* Madrid: Aliana. 527-529

Charlin J, Harnández Llosas I (2016) Morfometría Geométrica y Representciones Rupestres: Explorando las Aplicaciones de los Métodos Basados en Landmarks. *Arqueología* 22(1):103–125

Chartier, S.; Allarie, J.F. (2007) Power Estimation in Multivariate Analysis of Variance. *Tutorials in Quantitative Methods for Psychology*. 3(2):70-78 DOI: 10.20982/tqmp.03.2.p070

Chao, C.M.; Yu, Y.W.; Cheng, B.W.; Kuo, Y.L. (2014) Construction the Model on the Breast Cancer Survival Analysis use Support Vector Machine, Logistic Regression and Decision Tree, *Journal of Mdeical Systems*. 38(10) DOI: 10.1007/s10916-014-0106-1

Chevrier, B. (2012) *Les Assemblages à Pièces Bifaciales au Pléistocène Inférieur et Moyen Ancien en Afrique de l'Est et ay Proche-Orient*. PhD Thesis. Paris: Université Paris Ouest Nanterre la Défense.

Choi K.; Driwantoro, D. (2007) Shell tool use by early members of Homo erectus in Sangiran central java, Indonesia: cut mark evidence. *Journal of Archaeological Science.* 34:48–58

Chorowicz, J. (2005) The East African Rift System, *Journal of African Earth Sciences*. 43(1):379-410. DOI: 10.1016/j.jafrearsci.2005.07.019

Chollet, F. (2015) *Keras*. https://keras.io

Chollet, F.; Allaire, J.J. (2017) *Deep Learning with R*. Online E-Book Publisher: Manning Publications

Chollet, F. (2017) *Deep Learning with Python*. Online E-Book Publisher: Manning Publications.

Clark, J.D.; Haynes, C.V. (1970) An Elephant Butchery Site at Mwanganda's Village, Korongo, Malawi, and its Relevance for Paleolithic Archaeology, *World Archaeology.* 1:390-411

Clarke, D. (1973) Archaeology; The Loss of Inoccence, *Antiquity*. 47(185):7-18

Cobb, S.N.; O'Higgins, P. (2004) Hominins Do Not Share a Common Postnatal Facial Ontogenetic Shape Trajectory, *Journal of Experimental Zoology*. 302(B):302-321

Cobo-Sánchez, L.; Aramendi, J.; Domínguez-Rodrigo, M. (2014) Orientation Patterns of Wildebeest Bones on the Lake Masek Floodplain (Serengueti, Tanzania) and their Relevance to Interpret Anisotropy in the Olduvai Lacustrine Floodplain, *Quaternary International.* 322-323:277-284. DOI: 10.1016/j.quaint.2013.07.130

Cohen, J. (1988) *Statistical Power Analysis for Behavioural Sciences*. New York: Routledge.

Cook, J. (1986) Marked Human Bones from Gough's Cave, Somerset, *Proceedings of the University of Bristol Spelaeological Society*. 17(3):275-285

Coombs, W.T.; Algina, J.; Oltman, D.O. (1996) Univariate and Multivariate Omnibus Hypothesis Tests Selected to Control Type I Error Rates when Population Variances are not Necessarily Equal, *Review of Educational Research*. 66(2):137-179

Cooney, J,; Van Gelder, L. (2011) Child labour in the past: children as economic contributors and consumers. *Annual Society for the Study of Childhood in the Past*, University of Cambridge, Conference.

Coppens, Y. (1994) East Side Story: The Origin of Humankind, *Scientific American*. 88-95

Core, R Team (2018) A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/

Cortes, C.; Vapnik, V. (1995) Support-Vector Networks, *Machine Learning*. 20:273-297

Costa, A.G. (2010) A Geometric Morphometric Assessment of Shape in Bone and Stone Acheulean Bifaces from the Middle Pleistocene site of Castel di Guido, Latlum, Italy. In: S.J. Lycett, and P.R. Chauhan (Eds.) *New Perspectives on Old Stonces: Analytical Approaches to Palaeolithic Technologies*. New York: Springer. 23-42

Costamagno, S.; Liliane, M.; Cédric, B.; Bernard, V.; Bruno, M. (2006) Les Pradelles (Marillac-le-Franc, France): A Mousterian Reindeer Hunting Camp? *Journal of Anthropological Archaeology*. 25:466-484

Courtenay, L.A. (2017) *Tafonomía Aplicada al Estudio de las Marcas de Corte*. Bachelor's Dissertation. Madrid: Universidad Complutense de Madrid.

Courtenay, L.A.; Yravedra, J.; Maté-González, M.Á.; Aramendi, J.; González-Aguilera, D. (2017) 3D Analysis of Cut Marks using a New Geometric Morphometric Methodological Approach. *Journal of Archaeological and Anthropological Sciences*. DOI: 10.1007/s12520-017-0554-x

Courtenay, L.A.; Maté-González, M.Á.; Aramendi, J.; Yravedra, J.; González-Aguilera, D.; Domínguez-Rodrigo, M. (2018a). Testing Accuracy in 2D and 3D Geometric Morphometric Methods for Cut Mark Identification and Classification. *PeerJ*. 6:e5133, DOI: 10.7717/peerj.5133

Courtenay, L.A.; Yravedra, J.; Huguet, R.; Ollé, A.; Aramendi, J.; Maté-González, M.Á.; González-Aguilera, D. (2018b) New Taphonomic Advances in 3D Digital Microscopy: a Morphological Characterisation of Trampling Marks, *Quaternary International*. DOI: 10.1016/j.quaint.2018.12.019

Courtenay, L.A.; Yravedra, J.; Huguet, R.; Aramendi, J.; Maté-González, M.Á.; González-Aguilera, D.; Arriaza, M.C. (2019) Combining Machine Learning Algorithms and Geometric Morphometrics: a Study of Carnivore Tooth Marks, *Palaeogeography, Palaeoclimatology, Palaeoecology*. 522:28-29 DOI: 10.1016/j.palaeo.2019.03.007

Courtenay, L.A.; Aramendi, J.; Maté-González, M.Á.; Martín-Perea, D.M.; Yravedra, J.; Uribelarrea, D.; Baquedano, E.; González-Aguilera, D.; Domínguez-Rodrigo, M. (Under Review) Cut Marks and Raw Material Exploitation in the Lower Pleistocene Site of Bell's Korongo (BK, Olduvai Gorge, Tanzania): A Geometric Morphometric Analysis. *Quaternary International.*

Crevier, D. (1993) *The Tumultuous Search for Artificial Intelligence.* New York: Basic Books.

Crompton, R.H.; Growlett, J.A. (1993) Allometry and Multidimensional Form in Acheulian Bifaces from Kilombe, Kenya, *Journal of Human Evolution*. 25:175-199

Cruz-Uribe, K. (1991) Distinguishing Hyena from Hominid Bone Accumulations, *Journal of Field Archaeology*. 18(4):469-486.

Curtis, G.H.; Hay, R.L. (1972) Further Geological Studies and Potassium-Argon Dating at Olduvai Gorge and Ngorongoro Crater. In: W.W. Bishop, J.A. Miller (Eds.) *Calibration of Hominoid Evolution: Recent Advances in Isotopic and Other Dating Methods as Applicable to the Origin of Man.* Edinbugh: Scottish Academic Press. 289-301

Damodarasmy, S.; Raman, S. (1991) Texture Analysis using Computer Vision, *Computers in Industry*. 16:25-34

Darwin, C. (1859) On the Imperfection of the Geological Record, *The Origin of Species*. London: Wordsworth Editions. 9:279-311.

Dawkins, W.B. (1863) On a Hyaena Den at Wookey Hole, near Wells, *Quaternary Journal of the Geological Society*. 19(1-2):260-274. DOI: 10.1144/gsl.kgs.1863.019.01-02.27

Day, M.H. (1971) Postcranial Remains of Homo erectus from Bed IV, Olduvai Gorge, Tanzania, *Nature*. DOI: 10.1038/232383a0

Day, M.H. (1976) Hominid Postcranial Material from Bed I, Olduvai Gorge. In: G.L. Isaac and E.R. McCown (Eds.) *Human Origins: Louis Leaker and the East Africa Evidence*. California: WA Benjamin Inc. 363-374.

Day, M.H. (1986) *Guide to Fossil Man*. Chicago: Chicago University Press

Débenath, A.; Dibble, H.L. (1994) *Handbook of Paleolithic Typology. Vol 1, Lower and Middle Palaeolithic Europe*. Philadelphia: University of Pennsylvania

Dechter, R. (1986) Learning While Searching in Contraint-Satisfaction Problems. *National Conference on Artificial Intelligence*. 1:11-15

Defleur, A; Crégut-Bonnoure, E; Desclaux, E (1998) Première Mise en Évidence d'une Séquence éémienne à Restes Humains dans le Remplissage de la Baume Moula-Guercy (Soyons, Ardèche), *Comptes Rendus de l'Académie des Sciences*. 326(6):453-458.

Deino, A.L.; Scott, G.; Saylor, B.; Mulugeta, A.; Angelini, J.D.; Haile-Selassie, Y. (2010) 40Ar/39Ar Dating, Paleomagnetism and Tephrochemistry of Pliocene Strata of the Hominid-Bearing Woranso-Mille Area, West-Central Afar Rift, Ethiopia, *Journal of Human Evolution*. 58:111-126

Deino, A.L. (2012) 40Ar/39Ar Dating of Bed I, Olduvai Gorge, Tanzania, and the Chronology of Early Pleistocene Climate Change, *Journal of Human Evolution*. 63:251-273

Delany-Rivera, C.; Plummer, T.W.; Hodgson, J.A.; Forrest, F.; Hertel, F.; Oliver, J.S. (2009) Pits and Pitfalls: Taxonomic Variability and Petterning in Tooth Mark Dimensions, *Journal of Archaeological Science*. 36(11):2597-2608

Dempster, A.P.; Laird, N.M.; Rubin, D.B. (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society*. 39(1):1-38

Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L: (2009) ImageNet: a Large Scale Hierarchical Image Database, *Computer Vision and Pattern Recognition*. http://www.image-net.org/papers/imagenet_cvpr09.pdf

Deng, J.; Russakovsky, O.; Krause, J.; Bernstein, M.; Berg, A.; Fei-Fei, L. (2014) Scalable Multi-Label Annotation. *Conference of Human Factors in Computing.* http://ai.stanford.edu/~olga/papers/chi2014-MultiLabel.pdf

Deujeard, C.; Geraads, D.; Gallotti, R.; Lefèvre, D.; Mohib, A.; Raynal, J.P.; Hubin, J.J. (2016) Pleistocene Hominins as a Resource for Carnivores: A c. 500,000 Year Old Human Femur Bearing Tooth Marks in North Africa, *PLoS ONE*. DOI: 10.1371/journal.pone.0152284

Department of Education (U.K,) (2013) Science, *The National Curriculum in England, Keystages 1 and 2.* United Kingdom: Deparment of Education. [Online PDF] Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425601/PRIMARY_national_curriculum.pdf [Consulted: 24th October, 2018]. 144-172

Derech, N.; Tal, A.; Shimshoni, I. (2018) *Solving Archaeological Puzzles*. arXiv:1812.10553v1

Dibble, H.L.; Rezek, Z. (2009) Introducing a New Experimental Design for Controlled Studies of Flake Formation: Results for Exterior Platform Angle, Platform Depth, Angle of Blow, Velocity, and Force, *Journal of Archaeological Science*. 36(9):1945-1954.

Diedrich, C.G.; Žák, K. (2006) Prey Deposits and Den Sites of the Upper Pleistocene Hyena *Crocuta crocuta spelaea* (Goldfuss, 1823) in Horizontal and Vertical Caves of the Bohemian Karst (Czech Republic) *Bulletin of Geosciences*. 81(4):237-276

Diez-Martín, F. (2005) *El largo viaje*. Barcelon: Bellaterra.

Diez-Martín, F.; Sánchez, P.; Domínguez-Rodrigo, M.; Mabulla, A.; Barba, R. (2009) Were Olduvai Hominins making Butchering Tools or Battering Tools? Analysis of a Recently Excavated Lithic Assemblage from BK (Bed II, Olduvai Gorge, Tanzania), *Journal of Anthropological Archaeology*. 28:274-289

Diez-Martín, F.; Eren, M.I. (2012) The Early Acheulean in Africa: past paradigms, current ideas, and future directions. In: M. Dominguez-Rodrigo (Ed.): *Stone Tools and Fossil Bones. Debates in the Archaeology of Human Origins*. Cambridge: Cambridge University Press, Cambridge. 310–357.

Diez-Martín, F.; Sánchez-Yustos, P.; Uribelarrea, D.; Domínguez-Rodrigo, M.; Fraile-Márquez, C.; Obregón, R.A.; Díaz-Muñoz, I.; Mabulla, A.; Baquedano, E.; Pérez-González, A.; Bunn, H.T. (2014) New Archaeological and Geological Research at SHK Main Site (Bed II, Olduvai Gorge, Tanzania), *Quaternary International.* 322-323:107-128

Diez-Martín, F.; Sánchez-Yustos, P.; Uribelarrea, D.; Baquedano, E.; Mark, D.F.; Mabulla, A.; Fraile, C.; Duque, J.; Diaz, I.; Pérez-González, A.; Yravedra, J.; Egeland, C.P.; Organista, E.; Domínguez-Rodrigo, M. (2015) The Origin of the Acheulean: the 1.7 Million Year Old Site of FLK West, Olduvai Gorge (Tanzania), *Scientific Reports*. DOI: 10.1038/srep17839

Diez-Martín, F.; Fraile, C.; Uribelarrea, D.; Sánchez-Yustos, P.; Domínguez-Rodrigo, M.; Duque, J.; Díaz, I.; Francisco, S.; Yravedra, J.; Mabulla, A.; Baquedano, E. (2017) SHK Extension: a New Archaeological Window in the SHK Fluvial Landscape of Middle Bed II (Olduvai Gorge, Tanzania), *BOREAS*. DOI: 10.1111/bor.12246

Diez-Martín, F.; Wynn, T.; Sánchez-Yustos, P.; Duque, J.; Francisco, S.; Fraile, C. (2018) El Gran Bifaz de FLK-West y sus Implicaciones Cognitivas In: F. Diez-Martín (Ed.) *En África Hace 1,7 Millones de Años: El Origen del Achelense*. Madrid: Museo Arqueológico Regional. 92-103

Diez-Martín, F. (Ed.) (2018) *En África Hace 1,7 Millones de Años: El Origen del Achelense*. Madrid: Museo Arqueológico Regional.

Discamps, E.; Jaubert, J.; Bachellerie, F. (2011) Human Choices and Environmental Constraints: Deciphering the Variability of Large Game Procurement from Mousterian to Aurignacian Times (MIS 5-3) IN Southwestern France, *Quaternary Science Revies*. 30:2755-2775

Dixon, P.M. (2002) Ripley's K Function, *Encyclopedia of Environments*. 3:1796-1803

Doyle, A.C. (1890) The Sign of Four, *Lippincott's Magazine*. February Edition.

Domalain, M.; Bertin, A.; Daver, G. (2017) Was *Australopithecus afarensis* able to make the Lomekwian Stone Tools? Towards a Realistic Biomechanical Simulation of Hand Force Capability in Fossil Hominins and New Insights on the Role of the Fifth Digit. *Comptes Rendus Palevol*. 16:572-584

Domingos, P. (2015) *The Master Algorithm: How the Quest for the Ultimate Learning Machine will Remake our World.* New York: Basic Books.

Domínguez-Rodrigo, M. (1997) Meat-Eating by Early Hominids at the FLK 22 Zinjanthropus Site, Olduvai Gorge, Tanzania: an Experimental Approach Using Cut Mark Data, *Journal of Human Evolution*. 33(6):669-690.

Domínguez-Rodrigo, M. (1999) Flesh Availability and Bone Modification in Carcasses Consumed by Lions. *Palaeogeography, Palaeoclimatology, Palaeoecology*. 149(1-4):373-388.

Domínguez-Rodrigo, M.; Serrallonga, J.; Juan-Tresserras, J.; Alcalá, L.; Luque, L. (2001) Woodworking Activities by Early Humans: a Plant Residue Analysis on Acheulean Stone Tools from Peninj (Tanzania). *Journal of Human Evolution*. 40(4):289-299

Domínguez-Rodrigo, M.; Pickering, T.R. (2003) Early Hominid Hunting and Scavenging: a Zooarchaeological Review, *Evolutionary Anthropology*. 12:275-282

Domínguez-Rodrigo, M.; Piqueras, A. (2003) The Use of Tooth Pits to Identify Carnivore Taxa in Tooth-Marked Archaeofaunas and their Relevance to Reconstruct Hominid Carcass Processing Behaviours, *Journal of Archaeological Science*. 30(11):1385-1391

Domínguez-Rodrigo, M.; Pickering, T.R.; Semaw, S.; Rogers, M.J. (2005) Cutmarked Bones from Pliocene Archaeological Sites at Gona, Afar, Ethiopia: Implications for Function of the World's Oldest Stone Tools, *Journal of Human Evolution*. 48:109-121

Domínguez-Rodrigo, M.; Barba, R. (2006) New Estimates of Tooth Mark and Percussion Mark Frequencies at the FLK Zinj Site: The Carnivore-Hominid-Carnivore Hypothesis Falsified, Journal of Human Evolution. 50(2):170-194.

Domínguez-Rodrigo, M.; Barba, R. (2007) Five More Arguments to Invalidate the Passive Scavenging Version of the Carnivore-Hominid-Carnivore Model: A Reply to Blumenschine et al. (2007a), Journal of Human Evolution. 53(4):427-433.

Domínguez-Rodrigo, M.; Barba, R.; Egeland, C.P. (2007) *Deconstructing Olduvai*. The Netherlands: Springer

Domínguez-Rodrigo, M.; Yravedra, J. (2009) Why are Cut Mark Frequencies in Archaeofaunal Assemblages so Variable? A Multivariate Analysis, *Journal of Archaeological Science.* 36:884-894

Domínguez-Rodrigo, M.; Juana, S.; Galán, A.B.; Rodríguez, M. (2009a) A New Protocol to Differentiate Trampling Marks from Butchery Marks, *Journal of Archaeological Science*. 36(12):2643-2654 DOI: 10.1016/j.jas.2009.07.017

Domínguez-Rodrigo, M.; Mabulla, A; Bunn, H.T.; Barba, R.; Diez-Martín, F.; Egeland, C.P.; Espílez, E.; Egeland, A.; Yravedra, J.; Sánchez, P. (2009b) Unraveling Hominin Behavior at Another Anthropogenic Site from Olduvai Gorge (Tanzania): New Archaeological and Taphonomic Research at BK, Upped Bed II, *Journal of Human Evolution*. 57:260-283

Domínguez-Rodrigo, M.; Pickering, T.R. (2010) Un Estudio Tafonómico Multivariante de las Acumulaciones de Fauna de Hiénidos (Crocuta crocuta) y Félidos (Panthera pardus), *1º Reunión de Científicos Sobre Cubiles de Hiena (y otros grandes carnivoros) en los Yacimientos Arqueológicos de la Peninsula Ibérica*. 45-60.

Domínguez-Rodrigo, M.; Pickering, T.R.; Bunn, H.T. (2010) Configurational Approach to Identifying the Earliest Hominin Butchers, *PNAS*. DOI: 10.1073/pnas.1013711107

Domínguez-Rodrigo, M.; Fernández-López, S.R.; Alcalá, L. (2011a) How Can Taphonomy Be Defined in the XXI Century?, *Journal of Taphonomy*. 9(1):1-13

Domínguez-Rodrigo, M.; Pickering, T.R.; Bunn, H.T. (2011b) Reply to McPherron et al.: Doubting Dikika is about Data, not Paradigms. *PNAS* 108:E117-E117

Domínguez-Rodrigo, M.; Gidna, A.O.; Yravedra, J.; Musiba, C. (2012a) A Comparitive Neo-Taphonomic Study of Felids, Hyaenids and Canids: An Analogical Framework Based on Long Bone Modification Patterns, *Journal of Taphonomy*. 10(3)147-164

Domínguez-Rodrigo, M.; Pickering, T.R.; Bunn, H.T. (2012b) Experimental Study of Cut Marks made with Rocks Unmodified by Human Flaking and its Bearing on Claims of ~3.4-Million-Year-Old Butchery Evidence from Dikika, *Journal of Archaeological Sciences*. 39:205-214

Domínguez-Rodrigo, M.; Bunn, H.T.; Pickering, T.R.; Mabulla, A.Z.P.; Musiba, C.M.; Baquedano, E.; Ashley, G.M.; Diez-Martin, F.; Santonja, M.; Uribelarrea, D.; Barba, R.;

Yravedra, J.; Barboni, D.; Arriaza, C.; Gidna, A. (2012c) Autochtony and Orientation Patterns in Olduvai Bed I: a re-examination of the Status of Post-Depositional Biasing of Archaeological Assemblages from FLK-North (FLK-N), *Journal of Archaeological Science*. 39:2116-2127

Domínguez-Rodrigo, M; García-Pérez, A. (2013) Testing the Accuracy of Different A-Axis Types of Measuring the Orientation of Bones in the Archaeological and Palaeontological Record, *PLoS ONE* DOI: 10.1371/journal.pone.0068955

Domínguez-Rodrigo, M.; Pickering, T.R.; Baquedano, E.; Mabulla, A.; Mark, D.F. Musiba, C.; Bunn, H.T.; Uribelarrea, D.; Smith, V.; Diez-Martín, F.; Pérez-González, A.; Sánchez, P.; Santonja, M.; Barboni, D.; Gidna, A.; Ashley, G.; Yravedra, J.; Heaton, J.L.; Arriaza, M.C. (2013a) First Partial Skeleton of a 1.34 Million Year Old *Paranthropus boisei* from Bed II, Olduvai Gorge, Tanzania, *PLoS ONE*. DOI: 10.1371/journal.pone.0080347

Domínguez-Rodrigo, M.; Bunn, H.T.; Mabulla, A.Z.P.; Baquedano, E.; Uribelarrea, D.; Pérez-González, A.; Gidna, A.; Yravedra, J.; Diez-Martín, F.; Egeland, C.P.; Barba, R.; Arriaza, M.C.; Organista, E.; Ansón, M. (2013b) On Meat Eating and Human Evolution: A Taphonomic Analysis of BK4b (Upper Bed II, Olduvai Gorge, Tanzania), and its Bearing on Hominin Megafaunal Consumption, *Quaternary International*. DOI: 10.1016/j.quaint.2013.08.015

Domínguez-Rodrigo, M.; Diez-Martín, F.; Yravedra, J.; Barba, R.; Mabulla, A.; Baquedano, E.; Uribelarrea, D.; Sánchez, P.; Eren, M.I. (2014a) Study of the SHK Main Site Faunal Assemblage, Olduvai Gorge, Tanzania: Implications for Bed II Taphonomy, Paleoecology and Hominin Utilization of Megafauna, *Quaternary International*. 322-323:153-166

Domínguez-Rodrigo, M.; Uribelarrea, D.; Santonja, M.; Bunn, H.T.; García-Pérez, A.; Pérez-González, A.; Panera, J.; Rubio-Jara, S.; Mabulla, A.; Baquedano, E.; Yravedra, J.; Diez-Martín, F. (2014b) Autochtonous Anisotropy of Archaeological Materials by the Action of Water: Experimental and Archaeological Reassessment of the Orientation Patterns at the Olduvai Sites, *Journal of Archaeological Science*. 41:44-68

Domínguez-Rodrigo, M.; Yravedra, J.; Organista, E.; Gidna, A.; Fourvel, J.B.; Baquedano, E. (2015a) A New Methodological Approach to the Taphonomic Study of Paleontological and Archaeological Faunal Assemblages: a Preliminary Case Study from Olduvai Gorge (Tanzania), *Journal of Archaeological Science*. 59:35-53

Domínguez-Rodrigo, M.; Pickering, T.R.; Almécija, S.; Heaton, J.L.; Baquedano, E.; Mabulla, A.; Uribellarea, D. (2015b) Earliest Modern Human-Like Hand Bone from a new > 1.84 Million Year Old Site at Olduvai in Tanzania, *Nature Communications*. DOI: 10.1038/ncomms8987

Domínguez-Rodrigo, M.; Alcalá, L. (2016) 3.3 Million Year Old Stone Tools and Butchery Traces? More Evidence Needed. *PaleoAnthropology*. 2016:46-53

Domínguez-Rodrigo, M.; Cobo-Sánchez, L.; Yravedra, J.; Uribelarrea, D.; Arriaza, C.; Organista, E.; Baquedano, E. (2017a) Fluvial Spatial Taphonomy: a New Method for the Study of Post-Depositional Processes, *Journal of Archaeological and Anthropological Sciences*. DOI: 10.1007/s12520-017-0497-2

Domínguez-Rodrigo, M.; Saladié, P.; Cáceres, I.; Huguet, R.; Yravedra, J.; Rodríguez-Hidalgo, A.; Martín, R.; Pineda, A.; Marín, J.; Gené, C.; Aramendi, J.; Cobo-Sánchez, L. (2017b) Use and Abuse of Cut Mark Analyses: the Rorschach Effect, *Journal of Archaeological Science*. 86:14-23

Domínguez-Rodrigo, M. (2018) Successful Classification of Experimental Bone Surface Modifications (BSM) through Machine Learning Algorithms: a Solution to the

Controversial Use of BSM in Paleoanthropology? *Journal of Archaeological and Anthropological Sciences*. DOI: 10.1007/s12520-018-0684-9

Domínguez-Rodrigo, M.; Baquedano, E. (2018) Distinguishing Butchery Cut Marks from Crocodile Bite Marks through Machine Learning Methods. *Scientific Report*. DOI: 10.1030/s41598-018-24071-1

Domínguez-Rodrigo, M.; Cobo-Sánechez, L.; Aramendi, J.; Gidna, A. (2019) The Meta-Group Social Network for Early Humans: a Temporal-Spatial Assessment of Group Size at FLK-Zinj (Olduvai Gorge), *Journal of Human Evolution*. 127:54-66

Driver, J.C. (1990) Meat in Due Season: the Timing of Communal Hunts. In: L.B. Davis and B.O.K. Reeves (Eds.) *Hunters of the Recent Past*. London: Unwin Hyman. 11-33

Driver, J.C. (1995) Social Hunting and Multiple Predation. In: D.V. Campana (Ed.) *Before Farming: Hunter-Gatherer Society and Subsistence*. Philadelphia: University of Pensylvania Museum of Archaeology and Anthropology. 23-38

Dryden, I.L.; Mardia, K.V. (1998) *Statistical Shape Analysis*. New York: John Wiley and Sons.

Dryden, I.L. (2018) Shapes: Statistical Shape Analysis. https://cran.r-project.org/packages=shapes

Duches, R.; Nannini, N.; Fontana, A.; Boschin, F.; Crezzini, J.; Bernadini, F.; Tuniz, C.; Dalmeri, G. (2018) Archaeological Bone Injuries by Lithic Backed Projectiles: New Evidence on Bear Hunting from the Late Epigravettian Site of Cornafessa Rock Shelter (Italy). *Archaeological and Anthropological Sciences*. DOI: 10.1007/s12520-018-0674-y

Dunne, T.T.; Stone, M. (1993) Downdating the Moore Penrose Generalised Inverse for Cross-Validation of Centred Least Squares Prediction. *Journal of Royal Statistical Society B*. 55:369-375

Eaton, S.B.; Eaton, S.B. III; Cordain, L. (2002) Evolution, Diet and Health. In: P.S. Ungar and M.F. Teaford (Eds.) *Human Diet: Its Origin and Evolution*. Westport: Bergen and Garvey. 7-18

Ebinger, C.J. (2005) Continental Break Up: The East African Perspective, *Astronomical Geophysics*. 46:216-21

Echassoux, A. (2012) Comportements de Subsistance et Modifications Osseuses à l'aube de l'Acheuléen à Konso, Éthiopie, *L'Anthropologie*. 116:291-320

Eco, U.; Sebeok, T. (1989) *El Signo de los Tres; Dupin, Holmes, Peirce*. Barcelona: Editorial Lumen.

Efremov, I. (1940) Taphonomy: a New Branch of Palaeontology, *Pan American Geologist*. 74(2): 81-93.

Egeland, C.P.; Domínguez-Rodrigo, C. (2008) Taphonomic Perspectives on Hominid Site Use and Foraging Strategies during Bed II Times at Olduvai Gorge, Tanzania, *Journal of Human Evolution*. DOI: 10.1016/j.jhevol.2008.05.021

Egeland, C.P.; Domínguez-Rodrigo, M.; Pickering, T.R.; Menter, C.G.; Heaton, J.L. (2018) Hominin Skeletal Part Abundances and Claims of Deliberate Disposal of Corpses in the Middle Pleistocene, *PNAS*. DOI: 10.1073/pnas.1718678115

Ekins, S. (2016) The Next Era: Deep Learning in Pharmaceutical Research, *Pharmaceutical Research*. 33:2594-2603

Eldredge, N. (1989) Macroevolutionary Dynamics, In: J.A. Edler (Ed.) *Natural Selection in the Wild.* Princeton: Princeton University Press. 34.

Eliade, M. (1991) *Images and Symbols: Studies in Religious Symbolism*. New Jersey: Princeton University Press.

Esteban-Nadal, M.; Cáceres, I.; Fosse, P. (2010) Characterization of a Current Coprogenic Sample Originated by Canis lupus as a tool for Identifying a Taphonomic Agent, *Journal of Archaeological Science*. 37(12):2959-2970

Eulitz, M.; Reiss, G. (2015) 3D Reconstruction of SEM Images by Use of Optical Photogrammetry Software, *Journal of Structural Biology*. 191:190-196. DOI: 10.1016/j.jsb.2015.06.010

Evans, J. (1872) *The Ancient Stone Implements, Weapons and Ornaments of Great Britain*. London: Longmans, Green and Co.

Faith, J.T. (2007) Sources of Variation in Carnivore Tooth Mark Frequencies in a Modern Spotted Hyena (*Crocuta crocuta*) Den Assemblage, Amboseli Park, Kenya, *Journal of Archaeological Science*. 34(10):1601-1609

Faith, J.T.; Marean, C.W.; Behrensmeyer, A.K. (2007) Carnivore Competition, Bone Destruction, and Bone Density, *Journal of Archaeological Science*. 34(12):2025-2034

Faul, F.; Erdfelder, E.; Lang, A.G.; Buchner, A. (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behavior Research Methods*. 39(2):175-191.

Fawcett, T. (2006) An Introduction to ROC analysis. *Pattern Recognition Letters*. 27:861-874

Fernandes, K.; Chicco, D.; Cardoso, J.S.; Fernandes, J. (2018) Supervised Deep Learning Embeddings for the Prediction of Cervical Cancer Diagnosis, *PeerJ*. DOI: 10.7717/peerj-cs.154

Ferretti, M.P. (2007) Evolution of Bone-Cracking Adaptations in Hyaenids (Mammalia, Carnivore), *Swiss Journal of Geoscience*. 100:41-52. DOI: 10.1007/s00015-007-1212-6

Fernández-López, S.R. (1991) Taphonomic Concepts for a Theoretical Biochronology, *Revista Española de Paleontología*. 6(1):37-49

Fernández-López, S.R. (2000) *Temas de Tafonomía*. Madrid: Universidad Complutense de Madrid, Departamento de Paleoecologia.

Fernández-López, S.R.; Fernández-Jalvo, Y. (2002) The Limit Between Biostratinomy and Fossildiagenesis, *Current Topics on Taphonomy and Fossilization*. Conference Paper

Fernández-Jalvo, Y; Díez, JC; Cáceres, I; Rosell, J (1999) Human Cannibalism in the Early Pleistocene of Europe (Gran Dolina, Sierra de Atapuerca, Burgos, Spain), *Journal of Human Evolution*. 591-622.

Fernández-Jalvo, Y; Andrews, P (2003) Cannibalism in Britain: Taphonomy of the Creswellian (Pleistocene) Faunal and Human Remains from Gough's Cave (Somerset, England), *Bulletin of the Natural History Museum of London*. 58:59-81

Fernández-Jalvo, Y.; Marín Monfort, M.D. (2008) Experimental Taphonomy in Museums: Preperation Protocols for Skeletons and Fossil Vertebrates under the Scanning Electron Microscopy, *Geobios*. 41:157-181

Fernández-Jalvo, Y.; Andrews, P. (2016) *Atlas of Taphonomic Identifications*. Heidelberg: Springer.

Fenn, J.; Raskino, M. (2008) *Mastering the Hype Cycle: How to Choose the Right Innovation at the Right Time*. Boston: Harvard Business School Publishing.

Fernández-Marchena, J.L.; García-Argudo, G.; Pedergnana, A.; Valverde, I. (2016) *Líneas, Manchas y Cía. Pautas Metodológicas para una Adecuada Interpretación Funcional*. Conference: IX Jornadas de Jóvenes en Investigación Arqueológica. Santander 8-11 June, 2016. 241-250.

Feyerabend, P. (1993) *Against Method*. Vienna: Verso.

Fiorillo, A.R. (1984) An Experimental Study of Trampling: Implications for the Fossil Record. In: R. Bonnichsen and M.H. Sorg (Eds.) *Bone Modification.* Maine: University of Maine Press. 73-98.

Fisher, J.W. (1995) Bone Surface Modification in Zooarchaeology, *Journal Archaeological Method and Theory*. 2(1):7-68

Foley, C.; Foley, L.; Lobora, A.; Luca, D.; Msuha, M.; Davenport, T.R.B.; Durant, S. (2014) *Larger Mammals of Tanzania*. New Jersey: Princeton University Press.

Forbis, R.G. (1978) Some Facets of Communal Hunting, *Plains Anthropology*. 23:3-8

Fosse, P. (1996) La Grotte nº1 de Lunel-Viel (Hérault, France): Repaire d'Hyènes du Pléistocène Moyen. Etude Taphonomique du Matériel Osseux, *Paleo*. 8:47-79

Fosse, P. (1997) Variabilité des Assemblages Osseux Créés par l'Hyène des Cavernes, *Paleo*. 9:15-54

Franco, N.V.; Cardillo, M.; Borrero, L.A. (2005) Una Primera Aproximación de la Variabilidad Presente en las Puntas de Proyectil Denominadas "Bird IV", *Werken*. 6:81-95

Franco, N.V.; Castro, A.; Cardillo, M.; Charlin, J. (2009) La importancia de las variables morfológicas, métricas y de microdesgaste para evaluar las diferencias en diseños de puntas de proyectil bifaciales pedunculadas: un ejemplo del sur de Patagonia continental. *Magallania* 37(1):99-112

Frayer, D; Orschiedt, J; Cook, J; Russel, MD; Radovcic, J (2006) Kaprina 3: Cut Marks and Ritual Behaviour, *Periodicum Biologorum*. 108(4):519-524.

Freidline, S.E., Gunz, P., Harvati, K. and Hublin, J.-J. (2012a) A comprehensive morphometric analysis of the frontal and zygomatic bone of the Zuttiyeh fossil from Israel. *Journal of Human Evolution.* 62:225-241.

Freidline, S.E., Gunz, P., Harvati, K. and Hublin, J.J. (2013) Evaluating developmental shape changes in Homo antecessor subadult facial morphology. *Journal of Human Evolution* 65:404-423.

Frison, G.C. (2004) *Survival by Hunting. Prehistoric Human Predators and Animal Prey*. California: University of California Press.

Frost, S.R.; Marcus, L.F.; Bookstein, F.L.; Reddy, D.P.; Delson, E. (2003) Cranial Allometry, Phylogeography and Systematics of Large Bodied Papionins (Primates: Cercopithecinae) Inferred From Geometric Morphometric Analysis of Landmark Data. *The Anatomical Record Part A*. 275(A):1048-1072.

Fuentes-Sánchez, D.; Galindo-Pellicena, M.Á.; García-González, P.; Carretero, J.M.; Arsuage, J.L. (2017) Experimental Cut Marks Characterization using a Confocal Laser Profilometer. In: R. Alonso, J. Baena and D. Canales (Eds.) *Playing with the Time. Experimental Archaeology and the Study of the Past.* 171-176

Fullagar, R.; Furby, J.; Hardy, B. (1996) Residues on Stone Artifacts: State of a Scientific Art. *Antiquity*. 70:740-745

Fyhrie, D. (2010) The Mechanical Properties of Bone. In: E. Orwoll, J. Bilezikian, D. Vandeschueren (Eds.) *Osteoporosis in Men*. Oregon: Elsevier. 51-67

Gabucio, M.J.; Cáceres, I.; Rodríguez-Hidalgo, A.; Rosell, J.; Saladié, P. (2014) A Wildcat (*Felis silvestris*) Butchered by Neanderthals in Level O of the Abric Romaní site (Capellades, Barcelona, Spain, *Quaternary International.* 326-327:307-318

Galán, A.B.; Rodríguez, M.; Juana, S.; Domínguez-Rodrigo, M. (2008) A New Experimental Study on Percussion Marks and Notches and their Bearing on the Interpretation of Hammerstone-Broken Faunal Assemblages, *Journal of Archaeological Science*. 36:776-784.

Galton, F. (1888) Co-Relations and their Measurement, Chiefly from Anthropometric Data. *Proceedings of the Royal Society of London*. 45:135-145. DOI: 10.1098/rspl.1888.0082

Galton, F. (1907) Vox Populi, *Nature*. 1949(75):450-451

Garcia, J.; Martínez, K.; Carbonell, E. (2011) Continuity of the First Human Occupation in the Iberian Peninsula: Closing the Archaeological Gap. *C.R. Palevol*. 10:279-284.

García-Medrano, P.; Ollé, A.; Ashton, N.; Roberts, M.B. (2018) The Mental Template in Handaxe Manufacture: New Insights into Acheulean Lithic Technological Behavior at Boxgrove, Sussex, UK, *Journal of Archaeological Method and Theory*. DOI: 10.1007/s108916-08-9376-0

Gauci, J.; Conti, E.; Liang, Y.; Virochsiri, K.; He, Y.; Kaden, Z.; Narayanan, V.; Ye, X.; Chen, Z. (2019) Horizon: Facebook's Open Source Applied Reinforcement Learning Platform, *Computing Research Repository*. arXiv: 1811.00260v3

Gaudzinski, S. (1995) Wallertheim Revisited: a Re-Analysis of the Fauna from the Middle Palaeolithic Site of Wallertheim (Rheinhessen/Germany), *Journal of Archaeological Science*. 22:51-66

Gaudzinski, S. (1996) On Bovid Assemblages and their Consequences for the Knowledge of Subsistence Patterns in the Middle Palaeolithic. *Processions of the Prehistoric Society*. 62:19-39

Gawehn, E.; Hiss, J.A.; Schneider, G. (2016) Deep Learning in Drug Discovery, *Molecular Informatics*. DOI: 10.1002/minf.201501008

Gentry, A.W. (1967) *Pelorovis oldowayensis* Reck, an Extinct Bovid from East Africa. *The Bulletin of the British Museum (Natural History*). 14(7):25-29

Gentry, A.W.; Gentry, A. (1978a) Fossil Bovidae (Mammalia) of Olduvai Gorge, Tanzania. Part I. *Bulletin of the British Museum of Natural History (Geology)*. 29(4):289-446

Gentry, A.W.; Gentry, A. (1978b) Fossil Bovidae (Mammalia) of Olduvai Gorge, Tanzania. Part II. *Bulletin of the British Museum of Natural History (Geology)*. 30(1):1-83

Gerber, S.; Neige, P.; Eble, G.J. (2007) Combining Ontogenetic and Evolutionary Scales of Morphological Disparity: A Study of Early Jurassic Ammonites, *Evolution and Development*. 9(5):472-482.

Geribàs, N.; Mosquera, M.; Vergès, J.M. (2010) What Novice Knappers have to Learn to Become Expert Stone Toolmakers. *Journal of Archaeological Science*. 37(11):2857-2870

Gidna, A.; Yravedra, J.; Domínguez-Rodrigo, M. (2013) A Cautionary Note on the Use of Captive Carnivores to Model Wild Predator Behavior: a Comparison of Bone Modification Patterns on Long Bones by Captive and Wild Lions, *Journal of Archaeological Science*. 6:341-351

Géron, A. (2017) *Hands-on Machine Learning with Scikit-Learn and TensorFlow*. Sebastopol: O'Reilly

Gifford-Gonzalez, D. (1977) *Observations of Modern Human Settlements as an Aid to Archaeological Interpretation.* PhD Thesis. California: University of Berkeley.

Gifford-Gonzalez, D. (1985) The Third Dimension in Site Structure: an Experiment in Trampling and Vertical Dispersal, *American Antiquity*. 50(4):803-818

Gifford-Gonzalez, D. (1991) Bones Are Not Enough: Analogues, Knowledge, and Interpretive Strategies in Zooarchaeology, *Journal of Anthropological Archaeology*. 10(3):215-254.

Goldstein, J.I.; Newbury, D.E.; Echlin, P.; Joy, D.C.; Lyman, C.E.; Lifshin, E.; Sawyer, L.; Michael, J.R. (2003) *Scanning Electron Microscopy and X-Ray Microanalysis*. New York: Kluwer Academic/Plenum Publishers.

Gomez-Uribe, C.A.; Hunt, N. (2015) The Netflix Recommender System: Algorithms, Business Value, and Innovation, *ACM Transactions on Management Information Systems*. 6(4) DOI: 10.1145/2843948

González-Aguilera, D.; López-Fernández, L.; Rodríguez-González, P.; Guerrero, D.; Hernandez-Lopez, D.; Remondino, F.; Menna, F.; Nocerino, E.; Toschi, I.; Ballabeni A.; Giani, M. (2016) Development of an All Purpose Free Photogrammetric Tool. *The International Archives of the Photogrammetry*: *Remote Sensing and Spatial Information Sciences Conference*.

González-Álvarez, D. (2011) Arqueología, Folklore y Comunidades Locales: Los Castros en el Medio Rural Asturiano, *Complutum*. 22:133-153

Goodall, C.R. (1991) Procrustes Methods in the Statistical Analysis of Shape, *Journal of the Royal Statistical Society. Series B*. 53(2):285-339

Goodall, C.R.; Mardia, K.V. (1993) Multivariate Aspects of Shape Theory, *The Annals of Statistics*. 21(2):848-866

Gower, J. (1975) Generalized Procrustes Analysis, *Psychometrika*. 40(1):33-51. DOI: 10.1007/BF02291475

Grayson, D.K. (1984) *Quantitative Zooarchaeology: Topics in the Analysis of Archaeological Faunas*. Cambridge: Academic Press Inc.

Green, W.D.K. (1996) The Thin Plate Spline and Images with Curving Features. IN: K.V. Mardia, C.A. Gill and I.L. Dryden (Eds.) *Image and Fusion and Shape Variabiligy*. Leeds: University of Leeds Press. 79-87.

Green, N.E.; Swiontkowski, M.F. (2000) Skeletal Trauma in Children. S.A.: Médica Panamericana

Greenfield, H.J. (1999) The Origins of Metallurgy: Distinguishing Stone from Metal Cut-Marks on Bones from Archaeological Sites. *Journal of Archaeological Science* 26:797-808.

Greenfield, H.J. (2008) Metallurgy in the Near East, A Zooarchaeological Perspective on the Origins of Metallurgy in the Near East: Analysis of Stone and Metal Cut Marks on Bone from Israel. In: H. Selin (ed.) *Encyclopedia of the History of Science, Technology, and Medicine in Non-Western Cultures*. Netherlands: Springer. 1639-1647.

Griamud-Hervé, D (1998) The Arago 21 and 47 Hominid Endocast. L'Anthropologie 102(1):21-34

Grosman, L., Smikt, O.; Smilansky, E. (2008) On the Application of 3-D Scanning Technology for the Documentation and Typology of Lithic Artifacts, *Journal of Archaeological Science*. 35:3101-3110

Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; Chen, T. (2018) Recent Advances in Convolutional Neural Networks, *Pattern Recognition.* 77:354-377

Guipert, G.; Lumley, M.A.; Lumley, H.; Mafart, B. (2004) Three Dimensional Imagery: a New Look at the Tautavel Man

Guipert, G., de Lumley, M.A.; de Lumley, H. (2013) Restauration Virtuelle d'Arago 21, *Compres Rendus Palevol*. 13:51-59.

Gümrükçu, M. (2017) *Assessing the Effects of Fluvial Abrasion on Bone Surface Modifications using High Resolution 3D Scanning*. Master's Thesis. Colorado: Colorado State University

Gümrükçu, M.; Pante, M.C. (2018) Assessing the Effects of Fluvial Abrasion on Bone Surface Modifications using High-Resolution 3-D Scanning. *Journal of Archaeological Science: Reports* 21:208-221.

Günther, F.; Fritsch, S. (2010) Neuralnet: Training of Neural Networks, *The R Journal*. 2(1):30-38

Gunz, P.; Mitteroecker, P.; Bookstein, F.L. (2004a) Semilandmarks in Three Dimensions. In: D.E. Slice (Ed.) *Modern Morphometrics in Physical Anthropology*. New York: Plenum Publishers. 73-98

Gunz, P.; Mitteroecker, P.; Bookstein, F.L.; Weber, G.W. (2004b) Computer aided reconstruction of incomplete human crania using statistical and geometrical estimation methods. *Enter the past: computer applications and quantitative methods in archeology*. BAR Int. Ser. 1227, 96e98.

Gunz, P. (2005) Statistical and Geometric Reconstruction of Hominid Crania – *Reconstructing Australopthecine Ontogeny*. PhD Thesis. Wien: University of Wien.

Gunz, P.; Mitteroecker, P.; Neubauer, S.; Weber, G.W.; Bookstein, F.L. (2009) Principles for the Virtual Reconstruction of Hominin Crania, *Journal of Human Evolution*. 57:48-62 DOI: 10.1016/j.jhevol.2009.04.004

Gunz, P.; Neubauer, S.; Maureille, B.; Hublin, J.J. (2011) Virtual Reconstruction of the Le Moustier 2 Newborn Skull, Implications for Neandertal Ontogeny, *Paleo*. 22:155-172

Gunz, P.; Bulygina, E. (2012) The Mousterian Child from Teshik-Tash is a Neanderthal: A Geometric Morphometric Study of the Frontal Bone, *American Journal of Physical Anthropology*. 149:355-379

Gunz, P., Neubauer, S., Golovanova, L., Doronichev, V., Maureille, B. and Hublin, J.J. (2012) A unique modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neanderthal newborn from Mezmaiskaya. *Journal of Human Evolution* 62(2):300-313.

Gurney, K. (1997) *An Introduction to Neural Networks*. London: UCL Press.

Güth, A. (2012) Usig 3D Scanning in the Investigation of Palaeolithic Engravings: First Results of a Pilot Study, *Journal of Archaeological Sicence*. 39(10):3105-3114

Hall, B. (2003) Descent with Modification: the Unity Underlying Homology and Homoplasy as Seen Through an Analysis of Development and Evolution. *Biology Reviews*. 78(3):409-433

Han, R.; Yang, Y.; Li, X.; Ouyang, D. (2018) Predicting Oral Disintegrating Tablet Formulations by Neural Network Techniques, *Asian Journal of Pharmaceutical Sciences*. 13:336-342

Harmand, S.; Lewis, J.E.; Feibel, C.S.; Lepre, C.J.; Prat, S.; Lenoble, A.; Boës, X.; Quinn, R.L.; Brenet, M.; Arroyo, A.; Taylor, N.; Clément, S.; Daver, G.; Brugal, J.P.; Leakey, L.; Mortlock, R.A.; Wright, J.D.; Lokorodi, S.; Kirwa, C.; Kent, D.V.; Roche, H. (2015) 3.3 Million Year Old Stone Tools from Lomekwi 3, West Turkana, Kenya, *Nature*. DOI: 10.1038/nature14464

Harris, J.M. (1974) Pleistocene Giraffidae (Mammalia, Artiodactyla) from East Rudolf, Kenya. In R.J.G. Savage and S.C. Coryndon (Eds.) *Fossil Vertebrates of Affrica*. London: Academic Press.

Harvati, K., Hublin, J-J. and Gunz, P. (2010) Evolution of middle-late Pleistocene human cranio-facial form: A 3-D approach. *Journal of Human Evolution.* 59:445-464.

Haslam, M. (2006) Potential Misidentification of in Situ Archaeological Residues: Starch and Conidia, *Journal of Archaeological Science*. 33:114-121.

Hastie, T.; Tibshirani, R. (2017) Package 'mda': Mixture and Flexible Discriminant Analysis, R Package Vesion 0.4-10 https://cran.r-project.org/web/packages/mda/mda.pdf

Hauk, D.W.W.; Anderson, S. (1984) A New Statistical Procedure for Testing Equivalence in Two-Group Comparative Biovariability Trials, *Journal of Pharmacokinetic Biopharmacy*. 12:83-91. DOI: 10.1007/BF01063612

Hay, R.L. (1976) *Geology of the Olduvai Gorge*. London: University of California Press.
Hay, R.L.; Leakey, M.D. (1982) Fossil Footprints of Laetoli, *Scientific American*. 246(2), 50-57.
Hay, R.L., Kyser, T. K. (2001) Chemical sedimentology and paleoenvironmental history of Lake Olduvai, a Pliocene lake in northern Tanzania. *Geological Society of America Bulletin* 113:1505-1521.
Hayden, B. (1981) Subsistence and Ecological Adaptations of Modern Hunter/Gatherers. In: R.S.O. Harding and G. Teleki (Eds.) *Omnivorous Primates: Gathering and Hunting in Human Evolution*. New York: Columbia University Press. 344-421
Haynes, G. (1980) Evidence of Carnivore Gnawing on Pleistocene and Recent Mammalian Bones, *Paleobiology*. 6(3):341-351
Haynes, G. (1983) A Guide for Differentiating Mammalian Carnivore Taxa Responsible for Gnaw Damage to Herbivore Limb Bones, Paleobiology. 9(2):164-172
He, K.; Zhang, X.; Ren, S.; Sun, J. (2015) Deep Residual Learning for Image Recognition, *Computer Vision Foundation*. arXiv: 1512.03385
Hearne, T. (1770) Joannis Lelandi Antiquarii de Rebus Britannicis Collectanea. London: Richardson. LXIII-LXVI
Heaton, J. (2012) *An Introduction to the Math of Neural Networks*. Chesterfield: Heaton Research.
Hegazy, O.; Soliman, O.S.; Salam, M.A. (2013) A Machine Learning Model for Stock Market Prediction, *International Journal of Computer Science and Telecommunications*. 4(12):17-23
Henrique, B.M.; Sobreiro, V.A.; Kimura, H. (2018) Stock Price Prediction using Support Vector Regression on Daily and up to the Minute Prices, *Journal of Finance and Data Science*. 4:183-201
Heredia, B.; Khoshgoftaar, T.M.; Fazelpour, A.; Dittman, D.J. (2015) Building an Effective Classification Model for Breast Cancer Response data, *Information Re-Use and Integration*. DOI: 10.1109/IRI.2015.46
Hervé, M (2018) RVAideMemoire: Testing and Plotting Procedures for Biostatistics. https://cran.r-project.org/package=RVAideMemoire
Higuchi, S.; Chaminade, T.; Imamizu, H.; Kawato, M. (2009) Shared Neural Correlates for Language and Tool Use in Broca's Area, *Cognitive Neuroscience and Neuropsychology*. 20:1376-1381
Hillson, S. (2005) *Teeth*. Cambridge: Cambridge University Press. 2nd Edition
Hempel, J. (2018) Fei-Fei Li's Quest to Make AI Better for Humanity, *WIRED*. 26(5)
Hintze, J.L. (2007) *Descriptive Statistics, Means, Quality Control, and Design of Experiments*. Utah: NCSS
Hodder, I. (2008) 40 Years of Theoretical Engagement: a Conversation with Ian Hodder, *Norweigen Archaeological Review*. Interviewed by Hakan Karlsson and Bjornar Olsen. 41(1):26-42. DOI: 10.1080/00293650802181154
Holen, S.R.; Deméré, T.A.; Fisher, D.C.; Fullagar, R.; Paces, J.B.; Jefferson, G.T.; Beeton, J.M.; Cerutti, R.A.; Rountret, A.N.; Vescera, L.; Holen, K.A. (2017) A 130,000 Year Old Archaeological Site in Southern California, USA. *Nature* 544:479-483
Holtorf, C. (1998) The Life-Histories of Megaliths in Mecklenburg-Vorpommern (Germany), *World Archaeology*. 30:23-28
Horthorn, T., Hornick, K., Strobl, C., Zeileis, A. (2018) Package 'party': A Laboratory for Recursive Partytioning. R Package version 1.3-1. https://cran.r-project.org/web/packages/party/party.pdf

Hotelling, H. (1951) A Generalised T Test and Measure of Multivariate Dispersion. In J. Neyman (Ed.) *Processdings of the Second Berkley Symposium on Mathematical Statistics and Probability.* Berkley: University of California Press. 23-41

Houtson, D.C. (1979) The Adaptations of Scavengers. In: A.R.E. Sinclair and M.N. Griffiths (Eds.) *Serengeti: Dynamics of an Ecosystem.* Chicaco: University of Chicago Press. 263-286

Huerta, R.D. (2003) *The Parallel Search for Knowledge During the Age of Discovery.* Pensylvania: Bucknell University Press

Hughes, A.R. (1954) Hyaenas versus Australopithecines as Agents of Bone Accumulation, *American Journal of Physical Anthropology.* 12:467-486

Hunter, J.D. (2007) Matplotlib: a 2D Graphics Environment, *Computing in Science and Engineering*. 9:90-95. DOI: 10.1109/MCSE.2007.55

Hunter, J.; Dale, D.; Firing, E. (2019) *Matplotlib v.3.0.3*. https://matplotlib.org/

Hutton, J. (1794) *Experience Considered as the Cause of Knowledge, or as an Operation in which the Human Intellect is Made to Proceed, An Investigation of the Principles of Knowledge: and of the Progress of Reason, From Sense to Science and Philosophy*. London: A Strahan and T Cadell. 2(3):105-120

Iovita, R.; Tuvi-Arad, I.; Moncel, M.H.; Despriée, J.; Volnchet, P.; Bahain, J.T. (2017) High Handaxe Symmetry at the Beginning of the European Acheulian: the Data from La Noira (France) in Context, *PLoS ONE*. DOI: 10.1371/journal.pone.0177063

Isaac, G.L. (1971) The Diet of Early Man: Asepcts of Archaeological Evidence from Lower and Middle Pleistocene Sites in Africa, *World Archaeology*. 1:1-28

Isaac, G.L. (1977) *Olorgesailie: Archaeological Studies of a Middle Pleistocene Lake Basin in Kenya*. Chicago: University of Chicago Press.

Isaac, G.L. (1978) The Food Sharing Behaviour of Protohuman Hominids, *Scientific American.* 238:90-108.

Isaac, G.L. (1983) Bones in Contention: Competing Explanations for the Juxtaposition of Early Pleistocene Artifacts and Faunal Remains. In J. Clutton and C. Grigson (Eds.) *Animals and Archaeology*. Oxford: BAR International Series. 163:3-19

Jaubert, J.; Lorblanchet, M.; Laville, H.; Slott-Moller, R.; Turq, A.; Brugal, J.P. (1990) *Les Chasseurs d'Aurochs de la Borde: un site du Paleolithique Moyen (Livernon, Lot)*. Paris: Maison des Sciences de l'Homme.

Johanson, D.C.; Masao, F.T.; Eck, G.G.; White, T.D.; Walter, R.C.; Kimbel, W.H.; Asfaw, B.; Manega, P.; Ndessokia, P.; Suwa, G. (1987) New Partial Skeleton of *Homo habilis* from Olduvai Gorge, Tanzania, *Nature*. DOI: 10.1038/327205a0

Johnson, M. (2009) *Archaeological Theory: an Introduction*. London: Wiley

Jones, P. (1994) Results of Experimental Work in Relation to the Stone Industries of Olduvai Gorge. In: M.D. Leakey and D.A. Roe (Eds.): *Olduvai Gorge: Excavations in Beds III and IV, and the Masek Beds, 1968-1971.* Cambridge: Cambridge University Press.

de Juana, S.; Galán, A.B.; Domínguez-Rodrigo, M. (2010) Taphonomic Identification of Cut Marks made with Lithic Handaxes: an Experimental Study, *Journal of Archaeological Science*. 37:1841-1850 DOI: 10.1016/j.jas.2010.02.002

Jungers, W.L.; Falsetti, A.B.; Wall, C.E. (1995) Shape, Relative Size, and Size-Adjustments in Morphometrics, *Yearbook of Physical Anthropology*. 38:137-161.

Kalashnikov, D.; Irpan, A.; Pastor, P.; Ibarz, J.; Herzog, A:; Jang, E.; Quillen, D.; Holly, E.; Kalakrishnan, M.; Vanhoucke, V.; Levine, S. (2018) QT-OPT: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation, *Robot Learning* 2:1-23 arXiv: 1806.10293v3

Kaplan, A.; Haenlein, M (2018) Siri, Siri in my Hand, Who's the Fairest in the Land? On the Interpretations of Artificial Intelligence. *Business Horizons*, 62(1): 15-25

Keeley, L.H. (1980) *Experimental Determination of Stone Tool Uses: a Microwear Analysis*. Chicago: The University of Chicago Press

Keevil, T.L. (2018) *Inferring Early Stone Age Tool Technology and Raw Material from Cut Mark Micromorphology using high Resolution 3-D Scanning with Applications to Middle Bed II, Olduvai Gorge, Tanzania.* Master's Thesis. Colorado: Colorado State University.

Kendall, D. (1981) The Statistics of Shape. In V. Barnett (Ed.) *Interpreting Multivariate Data*. New York: Wiley. 75-80.

Kendall, D. (1984) Shape Manifolds, Procrustean Metric and Complex Projective Spaces. *Bulletin of the London Mathematical Society.* 16(2):81-121. DOI: 10.1112/blms/16.2.81

Kendall, D. (1989) A Survey of the Statistical Theory of Shape, *Statistical Science*. 4(2):87-89

Key, A.J.M.; Dunmore, C.J. (2015) The Evolution of the Hominin Thumb and the Influence Exerted by Non-Dominant Hand During Stone Tool Production, *Journal of Human Evolution*. 78:60-69

Key, A.J.M. (2016) Integrating Mechanical and Ergonomic Research within Functional and Morphological Analyses of Lithic Cutting Technology: Key Principles and Future Experimental Directions, *Ethnoarchaeology*. DOI: 10.1080/19442890.2016.1150626

Key, A.J.M.; Proffitt, T.; Stefani, E.; Lycett, S.J. (2016) Looking at Handaxes from Another Angle: Assessing the Ergonomic and Functional Importance of Edge form in Acheulean Bifaces, *Journal of Anthropological Archaeology*. 44:43-55.

Key, A.J.M.; Lycett, S.J. (2017a) Reassessing the Production of Handaxes versus Flakes from a Functional Perspective, *Journal of Archaeological Anthropological Sciences*. 9:737-753. DOI; 10.1007/s12520-015-0300-1

Key, A.J.M.; Lycett, S.J. (2017b) Influence of Handaxe Size and Shape on Cutting Efficiency: a Large Scale Experiment and Morphometric Analysis, *Journal of Archaeological Method and Theory*. 24:514-541. DOI: 10.1007/s10816-016-9276-0

Key, A.J.M.; Lycett, S.J. (2017c) Form and Function in the Lower Palaeolithic: History, Progress, and Continued Relevance, *Journal of Anthropological Sciences*. 95:1-42

Kim, B.J.; Kim K.S.; Lee, J.E.; Noh, D.Y.; Kim, S.W.; Jung, Y.S.; Park, M.Y.; Park, R.W. (2012) Development of Novel Breast Cancer Recurrence Prediction Model Using Support Vector Machine, *Journal of Breast Cancer*. 15(2):230-8

Kingdon, J.; Happold, D.; Hoffmann, M.; Butynski, T.; Happold, M.; Kalina, J. (2013) *Mammals of Africa. Volumes I to VI*. London: Bloomsbury.

Kingdon, J. (2015) *The Kingdon Field Guide to African Mammals*. London: Bloomsbury Publishing. 2nd Edition.

Kitagawa, K.; Krönneck, P.; Conard, N.J.; Münzel, S.C. (2012) Exploring Cave Use and Exploitation among Cave Bears, Carnivores and Hominins in the Swabian Jura, Germany. *Journal of Taphonomy* 10(3–4): 439–461.

Klein, R.G. (1976) The Mammalian Fauna of the Klasies River Mouth Sites, Southern Cape Province, South Africa, *The South African Archaeological Bulletin.* 31(123-124):75-98. DOI: 10.2307/3887730

Klein, R.G. (1989) *The Human Career: Human Biological and Cultural Origins.* Chicago: University of Chicago Press. 98-99

Klingenberg, C. (1998) Heterochrony and Allometry: The Analysis of Evolutionary Change in Ontogeny, *Biological Reviews*. 73:79-123

Klingenberg, C. (2008) Novelty and "Homology-Free" Morphometrics: What's in a Name? *Evolutionary Biology*. 35:186-190

Klingenberg, C. (2011) MorphoJ: an Integrated Software Package for Goemetric Morphometrics, *Molecular Ecology Resources*. 11:353-357

Klingenberg, C.; Marugán-Lobón, J. (2013) Evolutionary Covariation in Geometric Morphometric Data: Analyzing Integration, Modularity and Allometry in a Phylogenetic Context, *Systematic Biology*. 62(4):591-610 DOI: 10.1093/sysbio/syt025

Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, M.; Corlay, S.; Ivanov, P.; Avila, D.; Abdalla, S.; Willing, C.; Jupyter Development Team, *Positioning and Power in Academic Publishing.* DOI: 10.3233/978-1-61499-649-1-87.

Koopman, B.; Zuccon, G.; Nguyen, A.; Begheim, A.; Grayson, N. (2018) Extracting Cancer Mortality Statistics from Death Certificates: a Hybrid Machine Learning and Rule-Based Approach for Common and Rare Cancers, *Artificial Intelligence in Medicine.* 89:1-9

Koungoulos, L.; Faulkner, P.; Asmussen, B. (2018) Analysis of Pit and Score Tooth Mark Sizes from Bones Modified by Holocene Australian Terrestrial Fauna in Relation to Body Size, *Journal of Archaeological Science: Reports*. 20:271-283

Koepfli, K.P.; Jenks, S.M.; Eizirik, E.; Zahirpour, T.; Valkenburgh, B.V.; Wayne, R.K. (2006) Molecular Systematics of the Hyaenidae: Relationships of a Relictual Lineage Resolved by a Molecular Supermatrix, *Molecular Phyloenetics and Evolution*. 38:603-620

Kondo, O.; Dodo, Y.; Akazawa, T.; Muhesen, S. (2000) Estimation of Stature from the Skeletal Reconstruction of an Immature Neandertal from Dederiyeh Cave, Syria, *Journal of Human Evolution.* 38:457-473

Kondo, O.; Ishida, H.; Hanihara, T.; Wakebe, T.; Dodo, Y.; Akazawa, T. (2005) Cranial Ontogeny in Neandertal Children: Evidence from Neurocranium, Face and Mandible, *Current Trens in Dental Morphology Research.*

Korotcov, A.; Tkachenko, V.; Russo, D.P.; Ekins, S. (2017) Comparison of Deep Learning with Multiple Machine Learning Methods and Metrics using Diverse Drug Discovery Data Sets, *Molecular Pharmaceutics*. 14:4462-4475

Krasinski, K.E. (2016) Multivariate Evaluation of Criteria for Differentiating Cut Marks Created from Steel and Lithic Implements, *Quaternary International*. DOI: 10.1016/j.quaint.04.025

Kreutzer, L.A. (1992) Bison and Deer Bone Mineral Densities: Comparisons and Implications for the Interpretation of Archaeological Faunas, *Journal of Archaeological Science*. 19:271-294

Krizhevsky, A.; Sutskever, I.; Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems.* 25(2) DOI: 10.1145/3065386

Kruuk, H. (1972) *The Spotted Hyena: A Study of Predation and Social Behavior*. Chicago: University of Chicago Press.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kuhn, B.F. (2008) Examining Criteria for Identifying and Differentiating Fossil Faunal Assemblages Accumulated by Hyenas and Hominins using Extant Hyenid Accumulations, *International Journal of Osteoarchaeology*. 20(1):15-35 DOI: 10.1002/oa.996

Kuhn, M.; Johnson, K. (2013) *Applied Predictive Modelling*. Heidelberg: Springer

Kuhn, M. (2018) "Caret" Package: Classification and Regression Training. R Package Version 6.0-81 https://cran.r-project.org/web/packages/caret/caret.pdf

Kuhn, M., Weston, S., Culp, M., Coulter, N., Quinlan, R., 2018. C50: C5.0 Decision Trees and Rule-Based Models. R Package Version 0.1.2 https://cran.r-project.org/web/packages/C50/index.html

Lachenbruch, P.A. (1967) An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis, *Biometrics*. 23:639-645

Lagercrantz, S (1934) The Harpoon Down-Fall and its Distribution in Africa, *Anthropos*. 29(5):793-807

Lakens, D. (2017) Equivalence Tests: a Pratical Primer for T Tests, Correlations and Meta Analyses. *Society of Psychological and Personality Sciences*. 8(4):355-362

Lam, Y.M. (1992) Variability in the Behaviour of Spotted Hyaenas as Taphonomic Agents, *Journal of Archaeological Science*. 19(4):389-406

Lam, Y.M.; Chen, X.; Marean, C.W.; Frey, C.J. (1998) Bone Density and Long Bone Representation in Archaeological Faunas: Comparing Results from CT and Photon Densitometry, *Journal of Archaeological Science*. 25(6):559-570

Lam, Y.M.; Chen, X.; Pearson, O.M. (1999) Intertaxonomic Variability in Patterns of Bone Density and the Differential Representation of Bovid, Cervid, and Equid Elements in the Archaeological Record, *American Antiquity*. 64(2):343-362.

Lam, Y.M.; Pearson, O.M.; Marean, C.W.; Chen, X. (2003) Bone Density Studies in Zooarchaeology, *Journal of Archaeological Science*. 30(12):1701-1708

Lanier, J. (2018) *Ten Arguments for Deleting your Social Media Accounts Right Now*. New York: Henry Holt and Company.

Lansing, S.W.; Cooper, S.M.; Boydston, E.E.; Holekamp, K.E. (2009) Taphonomic and Zooarchaeological Implications of Spotted Hyena (*Crocuta crocuta*) Bone Accumulations in Kenya: a Modern Behavioral Ecological Approach. *Paleobiology*. 35:289-309

Lantz, B. (2013) *Machine Learning with R*. Birmingham: Packt Publishing.

Laplace, G. (1972) La Typologie Analytique et Structurale: Base Rationnelle d'Etude des Industries Lithiques et Osseuses, *Banques des Domnées Archéologiques*. Paris: CNRS. 932:91-143

Laplace, G.; Livache, M. (1972) Analytic Typology. Vitoria: Universidad del País Vasco

Lawley, D.N. (1938) A Generalisation of Fisher's Z-test, *Biometrika*. 30:180-187.

Lawrence, D. (1979) Taphonomy. *Encyclopaedia of Palaeontology*. Hutchinson & Ross, Sroudsburg. 793-799

Leakey, L.S.B. (1959) A New Fossil Skull from Olduvai, *Nature*. DOI: 10.1038/184491a0

Leakey, LS.B. (1960) Recent Discoveries at Olduvai Gorge, *Nature*. DOI: 10.1038/1881050a0

Leakey, L.S.B.; Leakey, M. (1964) Recent Discoveries of Fossil Hominids in Tanganyika, at Olduvai and Near Lake Natron, *Nature*. DOI: 10.1038/202005a0

Leakey, L.S.B.; Tobias, P.V.; Napier, J.R. (1964) A New Species of the Genus *Homo* from Olduvai Gorge, *Nature*. 4927:7-9

Leakey, L.S.B. (1965) *Olduvai Gorge, Volume 1: 1951-1961 Fauna and Background*. Cambridge: Cambridge University Press.

Leakey, M.D. (1965) A Descriptive List of the Named Localities in Olduvai Gorge. In: L.S.B. Leakey (Ed.) *Olduvai Gorge, Volume 1: 1951-1961 Fauna and Background*. Cambridge: Cambridge University Press. 101-107

Leakey, M.D. (1967) Preliminary survey of the cultural material from beds I and II, Olduvai Gorge, Tanzania. In: W. Bishop and J.D. Clark (Eds.) *Background to Evolution in Africa*. Chicago: University of Chicago Press. 417–446.

Leakey, M.D. (1971) *Olduvai Gorge, Volume 3: Excavations in Beds 1 and II, 1960-1963*. Cambridge: Cambridge University Press.

Leakey, M.D.; Harris, J.M. (1987) *Laetoli: a Pliocene Site in Northern Tanzania*. Oxford, Clarendon Press.

Leakey, M.D.; Roe, D.A. (1994) *Olduvai Gorge, Volume 5: Excavations in Beds III, IV and the Masek Beds 1968-1971*. Cambridge: Cambridge University Press.

LeCun, Y.; Cortes, C.; Burges, C.J.C. (1998a) *The MNIST Database of Handwritten Digits*. [Online] Available from: https://yann.lecun.com/exdb/mnist/ [Accessed on: 24/04/2019]

LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. (1998b) Gradient Based Learning Applied to Document Recongition, *Procedures of the IEEE*. 86(11):1-46

Lee, R.B.; DeVore, I. (1968) *Man the Hunter*. Chicago: Aldine.

Legg, S.; Hutter, M. (2007) A Collection of Definitions of Intelligence. *Frontiers in Artificial Intelligence and Applications*. 157:17-24 arXiv: 2007arxiv0706.3639L

Lei, L. (2018) Wavelet Neural Network Prediction Method of Stock Price Trend Based on Rough Set Attribute Reduction, *Applied Soft Computing*. 62:923-932

Lele, S.; Richtsmeier, J.T. (1991) Eculidean Distance Matrix Analysis: A Coordinate Free Approach for Comparing Biological Shapes using Landmark Data. *American Journal of Physical Anthropology*. 86:415-428. DOI: 10.1002/ajpa.1330860307

Lele, S.; Richtsmeier, J.T. (2001) *An Invaraint Approach to Statistical Analysis of Shapes*. USA: Chapman and Hall.

Lenoble, A.; Bertran, P. (2004) Fabric of Palaeolithic Levels: Methods and Implications for Site Formation Processes, *Journal of Archaeological Science*. 31:457-469

Leonard, W.R.; Robertson, M.L. (1992) Nutritional Requirements and Human Evolution: a Bioenergetics Model, *American Journal of Human Biology*. 4:179-195

Leonard, W.R.; Robertson, M.L. (1994) Evolutionary Perspectives on Human Nutrition: the Influence of Brain and Body Size on Diet and Metabolism, *American Journal of Human Biology*. 6:77-88

Lepre, C.J.; Roche, H.; Kent, D.V.; Harmand, S.; Quinn, R.L.; Brugal, J.P.; Texier, P.J.; Lenoble, A.; Feibel, C.S. (2011) An Earlier Origin for the Acheulian, *Nature*. DOI: 10.1038/nature10372

Leprinceince-Ringuet, D. (2018) Six Things you need to Know about the Future of Retail, *WIRED*. 17(58)

Lewis, J.E.; Hamand, S. (2016) An Earlier Origin for Stone Tool Making: Implications for Cognitive Evolution and the Transition to *Homo*, *Philosophical Transactions B.* DOI: 10.1098/rstb.2015.0233

Lieberman, D.E., McBratney, B.M. and Krovitz, G. (2002) The evolution and development of cranial form in Homo sapiens. *PNAS*. 99:1134-1139.

Linares-Matás, Gonzalo; Yravedra, José; Martín-Lerma, Ignacio; Aramendi, Julia; Courtenay, Lloyd A.; Maté-González, Miguel Ángel; Haber-Uriate, María; López-Martínez, Mariano; Walker, Michael J. (2017) Preliminary Tahponomical Assessment of the Macromammalian Zooarchaeological Assemblage at the Late Early Pleistocene Site of Cueva Negra del Estrecho del Río Quípar (Caravaca, Murcia, Spain), *European Society of Human Evolution (ESHE)*. Leiden: Stadsgehoorzaal.

Loizou, N; Rabbat, M.; Richtárik, P. (2019) Provably Accelerated Randomized Gossip Algorithms, *Facebook AI Research*. [Online] Available from: https://research.fb.com/publications/provably-accelerated-randomized-gossip-algorithms/ [Accessed 26/02/209]

Long, W.; Lu, Z.; Cui, L. (2019) Deep Learning Based Feature Engineering for Stock Price Movement Prediction, *Knowledge Based Systems*. 164:163-173

López-Cisneros, P.; Yravedra, J.; Álvarez-Alonso, D.; Linares-Matás, G. (2018) The Exploitation of Hunter Resources during the Magdalenian in the Cantabrian Region. Systematization of Butchery Processes at Coímbre Cave (Asturias, Spain), *Quaternary International*. DOI: 10.1016/j.quaint,2018.05.035

López-Polín, L.; Ollé, A.; Cáceres, I.; Carbonell, E.; Bermúdez de Castro, J.M. (2008) Pleistocene Human Remains and Conservation Treatments: the Case of a Mandible from Atapuerca (Spain). *Journal of Human Evolution*. 54(5):539-545.

López-Polín, L. (2012) Possible Inferences of Some Conservation Treatments with Subsequent Sutdies on Fossil Bones: a Conservator's Overview, *Quaternary International*. 275:120-127

Lozano, S.; Mateos, A.; Rodríguez, J. (2015) Exploring Paleo Food Webs in the European Early and Middle Pleistocene: a Network Analysis, *Quaternary International*. DOI: 10.1016/j.quaint.2015.10.068

Lyell, C. (1830) *Principles of Geology: Being an Attempt to Explain the Former Changes of the Earth's Surface, by Reference to Causes Now in Operation*. Wiltshire: Penguin Classics.

Lycett, S.J.; Cramon-Taubadel, N.; Foley, R.A. (2006) A Crossbeam Co-Ordinate Caliper for the Morphometric Analysis of Lithic Nuclei: a Description, Test and Empirical Examples of Application. Journal of Archaeological Science, 33:847-861

Lyman, R.L. (1987) Archaeofaunas and Butchery Sites: a Taphonomic Perspective. In: Schiffer, M. (Ed.) *Advances in Archaeological Method and Theory*. San Diego: Academic Press. 10:249-337

Lyman, R.L. (1994) *Vertebrate Taphonomy*. Cambridge: Manuals in Archaeology

Lyman, R.L. (2005) Analysing Cut Marks: Lessons from Artiodactyl Remains in the Northwestern United States, *Journal of Archaeological Science*. 32:1722-1732

Lyman, R.L. (2010) What Taphonomy Is, What it Isn't, and Why Taphonomists Chould Care about the Difference, *Journal of Tpaohnomy.* 8(1):1-16.

de Lumley, H.; de Lumley, M.A. (1971) Découverte de Restes Humaines Anté-Néandertaliens Datés au Début de Riss à la Caune d'Arago (Tautavel, Pyrénées-Orientales). *CR Acad. Sci Paris*. 272:1729-1742

de Lumley, H.; de Lumley, M.A. (1973) Preneanderthal Human Remains from Arago Cave in South-Eastern France. *Yerb Phys Anthrop* 17:162-168

Madurell-Malaperia, J.; Alba, D.M.; Minwer-Barakat, R.; Aurell-Garrido, J.; Moyà-Solà, S. (2012) Early Human Dispersals into the Iberian Peninsula: A Comment on Martínez et al. (2010) and Garcia et al. (2011), *Journal of Human Evolution*. 62(1):169-173

Mahmood, A.R.; Korenkevych, D.; Vasan, G.; Ma, W.; Bergstra, J. (2018) Benchmarking Reinforcement Learning Algorithms on Real-World Robots, *Proceedings of the Second Conference on Robot Learning.* arXiv: 1809.07731v1

Makysmilian, P. (1988) *Advanced Light Microscopy: Principles and Basic Properties*. Amsterdam: Elsevier

Makysmilian, P. (1989) *Advanced Light Microscopy: Specialised Methods.* Amsterdam: Elsevier

Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. (2016) Applications of Deep Learning in Biomedicine, *Molecular Pharmaceutics*. 13:1445-1454

Manega, P.C. (1993) Geochronology, Geochemistry and Isotopic Study of the Plio-Pleistocene Hominid Sites and the Ngorongoro Volcanic Highland in Northern Tanzania. PhD Thesis. Boulder: University of Colorado.

Mangasarian, O.L.; Street, W.N.; Wolberg, W.H. (1995) Breast Cancer Diagnosis and Prognosis via Linear Programming, *Operations Research*. 43(4):570-577.

Manríquez, G.; González-Bergás, F.; Carlos Salinas, J.; Espoueys, O. (2006) Deformación Intencional del Cráneo en Poblaciones Arqueológicas de Africa, Chile: Análysis Preliminar de Morfometría Geométrica con Uso de Radiografías Craneofacials. Revista de Antropología Chilena. 38(1):13-34.

Marín-Monfort, M.D.; Pesquero, M.D.; Fernández-Jalvo, Y. (2013) Compressive Marks from Gravel Substrate on Vertebrate Remains: A Preliminary Experimental Study. *Quaternary International*. DOI: 10.1016/j.quaint.2013.10.028

Marciani, G.; Arrighi, S.; Aureli, D.; Spagnolo, V.; Boscato, P.; Ronchitelli, A. (2018) Middle Palaeolithic Lithic Tools: Techno-Functional and Use-wear Analysis of Target Objects from SU 13 at the Oscurusciuto Rock Shelter, Southern Italy. *Journal of Lithic Studies* 5(2):1-30. DOI: 10.2218/jls.2745

Marean, C.W.; Spencer, L.M.; Blumenschine, R.J.; Capaldo, S. (1992) Captive Hyaena Bone Choice and Destruction, the Schlepp Effect and Olduvai Archaeofaunas, *Journal of Archaeological Science*. 19:101-121

Marean, C.W.; Bertino, L. (1994) Intrasite Spatial Analysis of Bone: Subtracting the Effect of Secondary Carnivore Consumers, *American Antiquity*. 59:748-768

Marean, C.W.; Frey, C.J. (1997) Animal Bones from Caves to Cities: Reverse Utility Curves as Methodological Artifacts, *American Antiquity*. 62(4):698-711.

Marshall, F.; Pilgrim, T. (1993) NISP vs MNI in Quantification of Body-Part Representation, *Society for American Archaeology*. 58(2):261-269

Martin, H. (1906) Presentation d'ossement de rene pertante des lesions d'origene humaine et animale. *Bulletin de la Societé Préhistorique Française*. 3:385-397.

Martin, H. (1907) Presentation d'ossements utilises de l'epoque Musterienne en Bourlon M. M. Giraux L. & Martin H. (1907). *Un os utilise presolutrean a propos de os utilices*. Nanterre: Société Préhistorique Française. 8-16

Martin, H. (1909). Desarticulation des quelques regions chez les rumiants et le cheval a l'epoque mousterienne. *Bulletin de la Societé Préhistorique Française* 7:303-310.

Martín, F.M. (2008) Bone Crunching Felids at the End of the Pleistocene in Fuego-Patafonia, Chile, *Journal of Taphonomy*. 6(3):337-372

Martínez, K.; García, J.; Carbonell, E.; Agustí, J.; Bahain, J.J.; Blain, H.A.; Burjachs, F.; Cáceres, I.; Duval, M.; Falguères, C.; Gómez, M.; Huguet, R. (2010) A New Lower Pleistocene Archaeological Site in Europe (Vallparadís, Barcelona, Spain), *PNAS*. 107:5762-5767

Martín-Viveros, J.I. (2016) *Microscopia digital, óptica y electrónica de barrido. Un enfoque complementario para el análisis de huellas de uso en industrias de sílex del Paleolítico medio. El nivel M del Abric Romaní (Capellades, Barcelona, España).* Master's Thesis, Tarragona: Universitat Rovira i Virgili.

Maslin, M.; Shutz, S.; Trauth, M. (2015) A synthesis of the theories and concepts of early human evolution. *Philosophical Transactions B.* 370(1663):20140064. DOI: 10.1098/rstb.2014.0064

Maté-González, M.A.; Yravedra, J.; González-Aguilera, D.; Palomeque-González, J.F.; Domínguez-Rodrigo, M. (2015) Microphotogrammetric characterization of cut marks on bones. *Journal of Archaeological Science* 62:128–142.

Maté-González, M.A.; Palomeque-González, J.F.; Yravedra, J.; González-Aguilera, D.; Domínguez-Rodrigo, M. (2016) Micro-photogrammetric and morphometric differentiation of cut marks on bones using metal knives, quartzite and flint flakes.

Journal of Archaeological and Anthropological Science. 10(4):805-816. DOI: 10.1007/s12520-016-0401-5

Maté-González, M.A.; Aramendi, J.; Yravedra, J.; González-Aguilera, D. (2017a) Statistical Comparison between Low-Cost Methods for 3D Characterization of Cut-Marks on Bones. *Remote Sensing*. 9(9):873. DOI: 10.3390/rs9090873

Maté-González, M.A.; Yravedra, J.; Martín-Perea, D.; Palomeque-González, J.; San-Juan-Blazquez, M.; Estaca-Gómez, V.; Uribelarrea, D.; Álvarez-Alonso, D.; Cuartero, F.; González-Aguilera D.; Domínguez-Rodrigo, M. (2017b) Flint and quartzite: Distinguishing raw material through bone cut marks. *Archaeometry*. 60(3):437-452. DOI: 10.1111/arcm.123275

Maté-González, M.A.; Aramendi, J.; Yravedra, J.; Blasco, R.; Rosell, J.; González-Aguilera, D.; Domínguez-Rodrigo, M. (2017c) Assessment of statistical agreement of three techniques for the study of cut marks: 3D Digital Microscope, Laser Scanning Confocal Microscopy and Micro-Photogrammetry. *Journal of Microscopy* 267(3):356-370. DOI: 10.1111/jmi.12575

Maté-González, M.Á. ; Courtenay, L.A. ; Aramendi, J. ; Yravedra, J. ; Domínguez-Rodrigo, M. ; Mora, R. ; González-Aguilera, D. (2019) Application of Geometric Morphometrics to the Analysis of Cut Mark Morphology on Different Bones of Different Sized Animals. Does Size Really Matter ? *Quaternary International.* DOI : 10.1016/j.quaint.2019.01.021

Matsumura, S. (2004) Orangutan Sociality, Tool Use, and Human Culture, *Current Anthropology*. 47(5):878

Maxmen, A. (2018) Deep Learning Spots Natural Selection at Work, *Nature Genetics*. 563:167

Mazloumian, A.; Eom, Y.H.; Helbing, D.; Lozano, S.; Fortunato, S. (2012) How Citation Boosts Promote Scientific Paradigm Shifts and Novel Prizes, *PLoS ONE*. 6(5):10.1371

McCorduck, P. (2004) *Machines Who Think*. Massachusetts: AK Peters Ltd.

McHenry, L.J.; Stanistreet, I.G. (2018) Tephrochronology of Bed II, Olduvai Gorge, Tanzania, and Placement of the Oldowan-Acheulean Transition, *Journal of Human Evolution*. 120:7-18. DOI: 10.1016/j.jhevol.2017.12.006

McKinney, W. (2010) Data Structures for Statistical Computing in Python, *Python in Science Conference*. 9:51-56

McNabb, J. (1989) Sticks and Stones: a Possible Experimental Solution to the Question of How the Clacton Spear Point was Made, *Proceedings of the Prehistoric Society*. 55.251-257

McPherron, S.P.; Alemseged, Z.; Marean, C.W.; Wynn, J.G.; Reed, D.; Geraads, D.; Bobe, R.; Béarat, H. (2010) Evidence for Stone-Tool-Assisted Consuption of Animal Tissues before 3.39 Million Years Ago at Dikika, Ethiopia, *Nature Letters*. 466:857-860. DOI: 10.1038/nature09248

McPherron, S.P.; Alemseged, Z.; Marean, C.; Wynn, J.G.; Reed, D.; Geraads, D.; Bobe, R.; Béarat, H. (2011) Tool Marked Bones from Before the Oldowan Change the Paradigm, *PNAS*. 108(21):E116. DOI: 10.1073/pnas.1101298108

Melis, A.P.; Schneider, A.C.; Tomasello, M. (2011) Chimpanzees, *Pan troglodytes*, Share Food in the Same Way after Collaborative and Individual Food Acquisition. *Animal Behaviour*. 82(3):485-493. DOI: 10.1016/j.anbehav.2011.05.024

Meltzer, D. (2006) *New Archaeological Investigations of a Classic Paleoindian Bison Kill*. Berkley: University of California Press.

Merino, J.M. (1994) *Tipologia Lítica*. San Sebastián: Sociedad de Ciencias Aranzadi

Merritt, S.R. (2012) Factors affecting Early Stone Age Cut Mark Cross-Sectional Size: Implications from Actualistic Butchery Trials, *Journal of Archaeological Science*. 39:2984-2994

Merritt, S.R. (2016) Cut Mark Cluster Geometry and Equifinality in Replicated Early Stone Age Butchery, *International Journal of Osteoarchaeology*. 26(4):585-598. DOI: 10.1002/oa.2448

Mersmann, O (2018) Mircobenchmark: Accurate Timing Functions. R Package Version 1.4-6 https://cran.r-project.org/package=microbenchmark

Merton, R.K. (1967) *On Theoretical Sociology: Five Essays, Old and New.* New York: The Free Press.

Merton, R.K. (1968) *Social Theory and Social Structure*. Massachusetts: Macmillan.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.C., Lin, C.C., 2018. e1071: Misc Functions of the Deparment of Statistics, Probability Theory Group. R Package version 1.7-0. https://cran.r-project.org/web/packages/e1071/index.html

Mevik, BH; Wehrens, R (2007) The pls Package: Principal Component and Partial Least Squares Regression in R, *Journal of Statistical Software*. 18(2):1-23

Mevik, B.H.; Wehrens, R., Liland, K.H., Hiemstra, P. (2018) Package 'pls': Partial Least Squares and Principal Component Regression. R Package version 2.7-0 https://cran.r-project.org/web/packages/pls/pls.pdf

Michel, P. (2005) Un Repaire Würmien d'Hyènes des Cavernes: la Grotte d'Unikoté (Iholdy, Pyrénées-Atlantiques, France). In: J.A. Lasheras, R. Montes (Eds.) *Actas de la Reunión Científica "Neandertales Cantábricos, Estado de la Cuestión*". Santander: Museo de Altamira. 165-177

Microsoft (2015) Visual Studio Code v. 1.32.3. https://code.visualstudio.com

Mills, M.G.L. (1990) *Kalahari Hyaenas: Comparative Behavioural Ecology of Two Species*. London: Unwin Hyman.

Milton, K. (1987) Primate Diets and Gut Morphology: Implications for Hominid Evolution. In: M. Harris and E.B. Ross (Eds.) *Food and Evolution: Toward a Theory of Human Food Habits*. Philadelphia: Temple University Press. 93-115

Mitchell, T.M. (1997) *Machine Learning.* New York: McGraw and Hill

Mitteroecker, P.; Gunz, P.; Berhard, M.; Schaefer, K.; Bookstein, F. (2004) Comparison of Cranial Ontogenetic Trajectories Among Great Apes and Humans, *Journal of Human Evolution*. 46(6):679-697. DOI: 10.1016/j.jhevol.2004.03.006

Mitteroecker, P.; Gunz, P.; Bookstein, F.L. (2005) Heterochrony and Geometric Morphometrics: A Comparison of Cranial Growth in Pan paniscus versus Pan troglodytes, *Evolution & Development*. 7(3):244-258. DOI: 10.1111/j.1525-142X.2005.05027

Mitteroecker, P. (2007) Evolutionary and Developmental Morphometrics of the Hominoid Cranium. PhD Thesis. Vienna: University of Vienna.

Mitteroecker, P.; Gunz, P. (2009) Advances in Geometric Morphometrics, *Evolutionary Biology*. 36:235-247. DOI: 10.1007/s11692-009-9055-x

Mitteroecker, P.; Bookstein, F.L. (2011) Linear Discrimination, Ordination, and the Visualization of Selection Gradients in Modern Morphometrics, *Evolutionary Biology*. 38:100-114. DOI: 10.1007/s11692-011-9109-8

Mitteroecker, P.; Gunz, P.; Neubauer, S.; Müller, G. (2012) How To Explore Morphological Integration in Human Evolution and Development?, *Evolutionary Biology*. 39:536-553. DOI: 10.1007/s11692-012-9178-3

Mitteroecker, P.; Gunz, P., Windhager, S.; Shaefer, K. (2013) A Brief Review of Shape, Form, and Allometry in Geometric Morphometrics, with Applications to Human Facial Morphology, *Hystrix, the Italian Journal of Mammology*. 24(1):59-66.

Moclán, A.; Domínguez-Rodrigo, M. (2018) An Experimental Study of the Patterned Nature of Anthropogenic Bone Breakage and its Impact on Bone Surface Modification

Frequencies, *Journal of Archaeological Science*. 96:1-13. DOI: 10.1016/j.jas.2018.05.007

Moclán, A.; Huguet, R.; Márquez, B.; Domínguez-Rodrigo, M.; Gómez-Miguelsanz, C.; Vergès, J.M.; Laplana, C.; Arsuaga, J.L.; Pérez-González, A.; Baquedano, E. (2018) Cut Marks made with Quartz Tools: an Experimental Framework for Understanding Cut Mark Morphology, and its Use at the Middle Palaeolithic Site of the Navalmaíllo Rock Shelter (Pinilla del Valle, Madrid, Spain), *Quaternary Interntional.* 493:1-18. DOI: 10.1016/j.quaint.2018.09.033

Moclán, A.; Domínguez-Rodrigo, M.; Yravedra, J. (2019) Classifying Bone Breakage Patterns: an Experimental Analysis of Fracture Planes to Discern Between Hominin and Carnivore Activity using Machine Learning (ML) Algorithms, *Anthropological and Archaeological Sciences*. DOI: 10.1007/s12520-019-00815-6

Monahan, C.M. (1996a) New Zooarchaeological Data from Bed II, Olduvai Gorge, Tanzania: Implications for Hominid Behavior in the Early Pleistocene, *Journal of Human Evolution*. 31:93-128

Monahan, C.M. (1996b) *Variability in the Foraging Behavior of Early Homo: a Taphonomic Perspective from Bed II, Olduvai Gorge, Tanzania*. PhD Thesis. Madison: University of Wisconsin.

Mondal, M.; Bertranpetit, J.; Lao, O. (2018) Approximate Bayesian Computation with Deep Learning Supports a Third Archaic Introgression in Asia and Oceania, *Nature Communications*. 10:246. DOI: 10.1038/s41467-018-08089-7

Monteiro, L.R.; Bordin, B.; Furtado dos Reis, S. (2000) Shape Distances, Shape Spaces and the Comparison of Morphometric Methods, *TREE*. 15(6):217-220

Montelius, O. (1888) *The Civilization of Sweden in Heathen Times*. London: Macmillan.

Moretti, E.; Arrighi, S.; Boschin, F.; Crezzini, J.; Aureli, D.; Ronchitelli, A. (2015) Using 3D Microscopy to Analyze Experimental Cut Marks on Animal Bones Produced with Different Stone Tools, *Ethnobiology Letters*. 6(2):14-22. DOI: 10.14237/ebl.6.1.2015.349

Moolayil, J. (2019) What's Next for Deep Learning Expertise, *Learn Keras for Deep Neural Networks*. Vancouver: Apress. 161-170

Moore, M.W.; Perston, Y. (2016) Experimental Insights into the Cognitive Significance of Early Stone Tools, *PLoS ONE*. DOI: 10.1371/journal.pone.0158803.

Müller, A. (2012) Kernel Approximations for Efficient SVMs (and other feature extraction methods). *Peekaboo*. [Online] Available at: https://peekaboo-vision.blogspot.com/2012/12/kernel-approximations-for-efficient.html?m=1

Münzel SC, Conard NJ (2004) Change and Continuity in Subsistence during the Middle and Upper Paleolithic in the Ach valley of Swabia (South-west Germany). *International Journal of Osteoarchaeology* 14: 225–243.

Muttart, M.V. (2017) *Taxonomic Distinctions in the 3D Micromorphology of Tooth Marks with Applcation to Feeding Traces from Middle Bed II, Olduvai Gorge, Tanzania*. Master's Thesis. Colorado: Colorado State University.

Myers, T.P.; Voorhies, M.R.; Corner, R.G. (1980) Spiral Fractures and Bone Pseudotools at Paleontological Sites, *American Antiquity*. 45(3):483-490

Nagabandi, A.; Kahn, G.; Fearing, R.S:; Levine, S. (2017) Neural Network Dynamics for Model Based Deep Reinforcement Learning with Model-Free Fine-Tuning, *Deep Reinforcement Learning Symposium*. arXiv: 1708.02596v2

Napier, J (1960) Fossil Hand Bones from Olduvai Gorge, *Nature*. 196:409-411. DOI: 10.1038/196409a0

Nelson, E.; Hall, J.; Randolph-Quinney, P.; Sinclair, A. (2017) Beyond size: the potential of a geometric morphometric analysis of shape and form for the assessment of sex in hand stencils in rock art. *Journal of Archaeolical Sciences*. 78:202–213

Nenadic, O.; Greenacre, M. (2007) Correspondence Analysis in R, with Two and Three Dimensional Graphics: The CA Package. *Journal of Statistical Software*. 20(3):1-13

Nick, L. (2015) The Unseen World: Reflections on Leeuwenhoek (1677), *Philosophical Trans Royal Society of London: Biological Sciences B*. 370:20140344. DOI: 10.1098/rstb.2014.0344

Nilsson, N. (1998) *Artificial Intelligence: A New Synthesis*. The Netherlands: Morgan Kauffman

Nilsson, P.J. (2000) *An Actualistic Butchery Study in South Africa and its Implications for Reconstructing Hominid Strategies of Carcass Acquisition and Butchery in the Upper Pleistocene and Plio-Pleistocene*. PhD Thesis. Cape Town: Cape Town University.

Njau, J. (2006) *The Relevance of Crocodiles to Oldowan Hominin Paleoecology at Olduvai Gorge, Tanzania*. PHD Thesis. New Jersey: Rutger State University.

Njau, J.; Blumenschine, R.J. (2006) A Diagnosis of Crocodile Feeding Traces on Larger Mammal Bone, with Fossil Examples from the Plio-Pleistocene Olduvai Basin, Tanzania, *Journal of Human Evolution*. 50(2):142-162

Njau, J.; Gilbert, H. (2016) Standardizing Terms for Crocodile-Induced Bite Marks on Bone Surfaces in Light of the Frequent Bone Modification Equifinality found to Result from Crocodile Feeding Behaviour, Stone Tool Modification, and Trampling, *FOROST 3*. DOI: 10.13140/RG.2.2.15139.66087

Nolle, T.; Luettgen, S.; Seeliger, A.; Mühlhäuser, M. (2018) Analyzing Business Process Anomalies Using Autoencoders, *Machine Learning*. 107:1875-1893 DOI: 10.1007/s10994-018-5702-8

Nowell, A.; Walker, C.; Cordova, C.E.; Ames, C.J.H.; Pokines, J.T.; Stueber, D.; DeWitt, R.; al-Souliman, A.S.S. (2016) Middle Pleistocene Subsistence in the Azraq Oasis, Jordan: Protein Residue and Other Proxies, *Journal of Archaeological Science*. 73:36-44

Öhman, C.; Zwierzak, I.; Beleani, M.; Viceconti, M. (2012) Human Bone Hardness Seems to Depend on Tissue Type but not on Anatomical Site in the Long Bones of an Old Subject, *Journal of Engineering in Medicine*. 227(2):200-206

O'Higgins, P.; Johnson, D. (1988) The Quantitative Description and Comparison of Biological Forms, *Critical Reviews in Anatomical Sciences*. 1:149-170

O'Higgins, P.; Jones, N. (2006) Morphologika 2 v2.5. York: University of York

Oliver, J.S. (1984) Analogues and Site Context: Bone Damages from Shield Trap Cave (24CB91), Carbon County, Montana, USA. In: R. Bonischen and M.H. Sorg (Eds.) *Bone Modification*. Maine: University of Maine Press. 61-72

Oliphant, T. (2006) *Numpy*. USA: Trelgol Publishing. http://www.numpy.org

Ollé, A. (2003) *Variabilitat i patrons funcionals en els sistmes tècnics de mode 2. Anàlisi de les deformaicons d'ús en els conjunts lítics del Riparo Estertno de Grotta Paglicci (RignanoGarganico, Foggia), Áridos (Arganda, Madrid) i Galería-TN (Sierra de Atapuerca, Burgos)*. PhD Thesis. Tarragona: Universitat Rovira i Virgili

Olsen, S.L. (1988) The Identification of Stone and Metal Tool Marks on Bone Artefacts, *Scanning Electron Microscopy in Archaeology*. BAR Int Ser. 452:337-360

Olsen, S.L.; Shipman, P. (1988) Surface Modification on Bone: Trampling Versus Butchery, *Journal of Archaeological Science*. 15(5):535-553.

Organista, E.; Domínguez-Rodrigo, M.; Egeland, C.P.; Uribelarrea, D., Mabulla, A.; Baquedano, E. (2016) Did *Homo erectus* Kill a *Pelorovis* Herd at BK (Olduvai Gorge)? A Taphonomic Study of BK5, *Archaeological and Anthropological Sciences*. 8:601-624

Organista, E.; Domínguez-Rodrigo, M.; Yravedra, J.; Uribelarrea, D.; Carmen Arriaza, M.; Ortega, M.C.; Mabulla, A.; Gidna, A.; Baquedano, E. (2017) Biotic and Abiotic Processes Affecting the Formation of BK Level 4c (Bed II, Olduvai Gorge) and their Bearing on Hominin Behavior at the Site, *Palaeogeography, Palaeoclimatology, Palaeoecology*. 488:59-75

Organista, E. (2017) *Estudio Tafonómico de los Niveles Arqueológicos de Bell Korongo (BK) Garganta de Olduvai, Tanzania*. PhD Thesis. Madrid: Universidad Complutense de Madrid.

Orlikoff, E.R.; Keevil, T.L.; Pante, M.C. (2017a) Quantitiative Analysis of the Micromorphology of Trampling-Induced Abrasion and Stone Tool Cut Marks on Bone Surfaces. *Conference: Paleoanthropology Society Conference 2018*. DOI: 10.13140/RG.2.2.31360.94721

Otárolla-Castillo, E.; Torquato, M.; Hawkins, H.C.; James, E; Harries, J.A.; Marean, C.W.; McPherron, S.P.; Thompson, J.C. (2017b) Differentiating between Cutting Actions on Bone using 3D Geometric Morphometrics and Bayesian Analyses with Implications to Human Evolution. *Journal of Archaeological Science*. 89:56-67 DOI: 10.1016/j.jas.2017.10.004

The Oxford English Dictionary (2018) *The English Oxford Dictionary*. Oxford: Oxford University Press.

Oxnard, C.E. (1986) The Measurement of Form: Beyond Biometrics, *The Cleft Palate Journal Supplement*. 23:110-128

Oxnard, C.; O'Higgins, P. (2008) Biology Clearly Needs Morphometrics. Does Morphometrics Need Biology?, *Biological Theory*. 4(1):84-97

Pales, L; Lambert, C (1971) Mammifères du Quaternaire, *Atlas Ostéologique pour server à l'Identification des mammiferes du Quaternaire*. Paris: Centre National de la Recherche Scientifique. 2:14-15

Palomeque-González, J.F.; Maté-González, M.Á.; Yravedra, J.; Juan-Blazquez, M.S.; Vargas, E.G.; Martín-Perea, D.M.; Estaca-Gómez, V.; González-Aguilera, D.; Domínguez-Rodrigom M. (2017) *Pandora*: A New Morphometric and Statistical Software for Analysing and Distinguishing Cut Marks on Bones, *Journal of Archaeological Science: Reports*. 13:60-66.

Pante, M.C.; Blumenschine, R.J. (2010) Fluvial Transport of Bovid Long Bones Fragmented by the Feeding Activities of Hominins and Carnivores, *Journal of Archaeological Science*. 37:846-854

Pante, M.C.; Blumenschine, R.J.; Capaldo, S.; Scott, R.S. (2012) Validation of Bone Surface Modification for Inferring Hominin and Carnivore Feeding Interactions with Reaplication to FLK 22, Olduvai Gorge, Tanzania, *Journal of Human Evolution*. 63:395-407

Pante, M.C.; Muttart, M.V.; Keevil, T.L.; Blumenschine, R.J.; Njau, J.K.; Merritt, S.R. (2017a) A new high-resolution 3-D quantitative method for identifying bone surface modifications with implications for the Early Stone Age archaeological record. *Journal of Human Evolution* 102:1-11.

Pante, M.C.; Njau, J.K.; Hensley-Marschand, B.; Keevil, T.L.; Martín-Ramos, C.; Peters, R.F.; Torre, I. (2017b) The Carnivorous Feeding Behavior of Early *Homo* at HWK EE, Bed II, Olduvai Gorge, Tanzania, *Journal of Human Evolution*. 120:215-235. DOI: 10.1016/j.jhevol.2017.06.005

Pappu, S.; Gunnell, Y.; Akhilesh, K.; Braucher, R.; Taieb, M.; Demory, F.; Thoveny, N. (2011) Early Pleistocene Presence of Acheulian Hominins in South India, *Science*. 331(6024):1596-1599. DOI: 10.1126/science.1200183

Pappu, S.; Akhilesh, K. (2019) Tools, Trials and Time: Debating Acheulian Group Size at Attirampakkam, India, *Journal of Human Evolution*. 130:109-125

Parrinder, G. (1971) Ancient Egypt, *World Religions from Ancient History to the Present*. Bicester: Facts on File Publications. 135-145

Parsons, T. (Ed.) (1968) *Knowledge and Society – American Sociology*. New York: Basic Books.

Parsons, T. (1975) *Social Systems and the Evolution of Action Theory*. New York: the Free Press.

Patel, A.A. (2019) *Hands-on Unsupervised Learning Using Python.* Sebastopol: O'Reilly.

Paterson, J.; Gibson, A. (2017) *Deep Learning*. Sebastopol: O'Reilly.

Pedregosa, F.; Varoquax, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. (2011) Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*. 12:2825-2830

Perez, S.I.; Bernal, V.; Gonzalez, P.N. (2006) Differences between Sliding Semi-Landmark Methods in Geometric Morphometrics, with an Application to Human Craniofacial and Dental Variation. 208(6):769-784. DOI: 10.1111/j.1469-7580.2006.00576.x

Peterhans, J.C.K.; Singer, R. (2006) Taphonomy of a Lair Near the Peers (or Skildegat) Cave in Fish Hoek, Western Cape Province, South Africa, *South African Archaeological Bulletin.* 61(183):2-18

Petö, R.; Elekes, F.; Oláh, K.; Király, I. (2018) Learning How to Use a Tool: Mutually Exclusive Tool-Function Mappings are Selectively Acquired from Linguistic In-group Models, *Journal of Experimental Child Psychology*. 171:99-112. DOI: 10.1016/j.jecp.2018.02.007

Pickering, T.R. (2002) Reconsideration of Criteria for Differentiating Faunal Assemblages Accumulated by Hyenas and Hominids, *International Journal of Osteoarchaeology*. 12:127-141. DOI: 10.1002/oa.594

Pickering, T.R.; Marena, C.W.; Domínuez-Rodrigo, M. (2003) Importance of Limb Bone Shaft Fragments in Zooarchaeology: A Response to "On *in situ* Attrition and Vertebrate Body Part Profiles" (2002), by M. C. Stiner, *Journal of Archaeological Science*. 30(11):1469-1482

Pickering, T.R.; Egeland, C.P. (2006) Experimental Patterns of Hammerstone Percussion Damage on Bones: Implications for Inferences of Carcass Processing by Humans, *Journal of Archaeological Science*. 33(4):459-469.

Pickering, T.R.; Heaton, J.L.; Zwodeski, S.E.; Kuman, K. (2011) Taphonomy of Bones from Baboons Killed and Eaten by Wild Leopards in Mapungubwe National Park, South Africa, *Journal of Taphonomy*. 9(2):117-159

Pineda, A.; Saladié, P.; Vergés. J.M.; Huguet, R.; Cáceres, I.; Vallverdú, J. (2014) Trampling versus Cut Marks on Chemically Altered Surfaces: an Experimental Approach and Archaeological Application at the Barranc de la Boella Site (la Canonja, Tarraona, Spain). *Journal of Archaeological Science.* 50:84-93.

Pineda, A.; Saladié, P. (2018) The Middle Pleistocene Site of Torralbo (Soria, Spain): A Taphonomic View of the Marquis of Cerralbo and Howell Faunal Collections, *Archaeological and Anthropological Sciences*. 11(6):2539-2556. DOI: 10.1007/s12520-018-0686-7

Pineda, A.; Cáceres, I.; Saladié, P.; Huguet, R.; Rosas, A.; Vallverdú, J. (2019) Tumbling Effects on Bone Surface Modificaitions (BSM): An Experimental Application on Archaeological

Deposits from the Barranc de la Boelle Site (Tarragona, Spain), *Journal of Archaeological Science*. 102:35-47. DOI: 10.1016/j.jas.2018.12.01

Piperno, D.P. (1988) *Phytolith Analysis – An Archaeological and Geological Perspective*. New York: Academic Press.

Playfair, J. (1802) Figure of the Earth, Illustrations of the Huttonian Theory of the Earth. Edinburgh: William Creech. 488-509.

Pobiner, B.I.; Braun, D.R. (2005) Applying Actualism: Considerations for Future Research. *Jouran of Taphonomy*. 3(2):57-65

Pokines, J.T.; Kerbis-Peterhans, J.C. (2007) Spotted Hyena (*Crocuta crocuta*) Den Use and Taphonomy in the Masai Mara National Reserve, Kenya, *Journal of Archaeological Sciences*. 34:1914-1931

Popper, K. (1935) *The Logic of Scientific Discovery*. London: Routledge.

Popper, K. (1963) *Science: Conjectures and Refutations – The Growth of Scientific Knowledge*. London: Routledge

Potter, S.L. (2005) The Physics of Cut Marks, *Journal of Taphonomy* 3(2):91-106

Potts, R.; Shipman, P. (1981) Cutmarks Made by Stone Tools on Bones from Olduvai Gorge, Tanzania, *Nature*. 291:577-1981

Potts, R. (1982) *Lower Pleistocene Site Formation and Hominid Activities at Olduvai Gorge, Tanzania*. PhD Thesis. Harvard: Harvard University.

Potts, R. (1988) *Early Hominid Activities at Olduvai*. New York: Aldine and Gruyter.

Potts, R. (1996) Evolution and Climate Variability, *Science*. 273:922-923.

Poole, D.; Mackworth, A.; Goebel, R. (1998) *Computational Intelligence: a Logical Approach*. Reading: Addison-Wesley.

Prendergast, M.E.; Domínguez-Rodrigo, M. (2008) Taphonomic Analyses of a Hyena Den and a Natural-Death Assemblage near Lake Eyasi (Tanzania), *Journal of Taphonomy*. 6(3-4):301-335

Putt S.S.; Wikeakumar S.; Franciscus R.G.; Spencer J.P. (2017) The Functional Brain Networks that Underlie Early Stone Age Tool Manufacture, *Natural Human Behaviour*. DOI: 10.1038/s41562-017-0102

Python Software Foundation (2018) Python Language v.2.7. http://www.python.org

Kolfschoten, T.V. (2017) Expression of Concern "Hominin Dispersals from the Jaramillo Subchron in Central and South-Western Europe: Untermassfeld (Germany) and Vallparadís (Spain)", *Quaternary International*. DOI: 10.1016/j.quaint.2013.03.005

Quinlan, J.R. (1992) *Learning with Continuous Classes*. Tasmania: Proceedings of the 5th Australian Joint Conference on Artificial Intelligence. 343-348.

Quinlan, J.R: (1996) Improved Use of Continuous Attributes in C4.5, *Journal of Artificial Intelligence Research*. 77-90.

Randles, C. (2017) *Jupyter Notebook*. https://jupyter.org

Rao, C.R. (1951) An Asymptotic Expansion of the Distribution of Wilks' Criterion, *Bulletin de l'Institut International Statistique.* 33:177-180.

Reck, H. (1933) *Olduvai: The Gorge of Primitive Man*. Leipzig: F.A. Brockhaus

Redford, D. (2002) *The Ancient Gods Speak: A Guide to Egyptian Religion*. Oxford: Oxford University Press.

Redmon, J.; Divvala, S.; Girschick, R.; Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection, *Computer Vision and Pattern Recognition*. arXiv: 1506.02640v5 https://pjreddie.com/yolo

Reinhart, A. (2015) *Statistics Done Wrong: The Woefully Complete Guide*. San Francisco: No Starch Press Inc.

Rendu, W.; Bourguignon, L.; Costamagno, S.; Meignen, L.; Soulier, M.C.; Armand, D.; Beuval, C.; David, F.; Griggo, C.; Jaubert, J.; Maureille, B.; Park, S.J. (2009) Mousterian Hunting Camps: Interdisciplinary Approach and Methodological Considerations. In: F. Bon, S. Costamagno, N. Valdeyron (Eds.) *Hunting Camps in Prehistory*. *Current Archaeological Approaches*. Toulouse: University of Toulouse II. 66-76

Rendu, W.; Costamagno, S.; Meignen, L.; Soulier, M.C. (2012) Monospecific Faunal Spectra in Mousterian Contexts: Implications for Social Behavior, *Quaternary International*. 247:50-58

Renne, P.R.; Swisher, C.C.; Deino, A.L.; Karner, D.B.; Owens, T.L.; DePaolo, D.J. (1998) Intercalibration of standards, absolute ages and uncertainties in 40Ar/39Ar dating. *Chemical Geology*. 145:117– 152.

Renne, P.R.; Balco, G.; Ludwig, K.R.; Mundil, R.; Min, K. (2011) Response to the comment by W.H. Schwarz et al. on "Joint determination of 40K decay constants and 40Ar/40K for the Fish Canyon sanidine standard, and improved accuracy for 40Ar/39Ar geochronology" by P.R. Renne et al., 2010. *Geochimica et Cosmochimica Acta*. 75:5097-5100.

Renfrew, C. (1973a) The Explanation of Cultural Change*: Models in Prehistory*. London: Duckworth

Renfrew, C. (1973b) *Social Archaeology*. Southampton: The Camelot Press.

Renfrew, C. (1989) Comments on Archaeology in the 1990s, *Norwegian Archaeological Review*. 22:33-41

Renfrew, C. (2014) Who Owns the Past? Whose Past? Who Tells the Story? Whose Story?, *The Future of Archaeology: The History of Archaeology, An Introduction.* London: Routledge. 245-246.

Renfrew, C.; Bahn, P. (2016) *Archaeology*. London: Thames & Hudson. 6th Edition.

Reynard, J.P. (2014) Trampling in Coastal Sites: an Experimental Study on the Effects of Shell on Bone in Coastal Sediment, *Quaternary International* 330:156-170

Rice, S.H. (1997) The Analysis of Ontogenetic Trajectories: When a Change in Size or Shape is not Heterochrony, *Proc. Natl. Acad. Sci*. 94:902-912.

Richter, A.N.; Khoshgoftaar, T.M. (2018) A Review of Statistical and Machine Learning Methods for Modelling Cancer Risk using Structured Clinical Data, *Artificial Intelligence in Medicine*. 90:1-14

Richtsmeier, J.T.; Paik, C.H.; Elfert, P.C.; Cole, T.M.; Dahlman, H.R. (1995) Precision, Repeatability and Validation of the Localization of Cranial Landmarks using Computed Tomography Scans, *Cleft Palate-Craniofacial Journal*. 32(3):217-227

Ripley, B.D. (1977) Modelling Spatial Patterns, *Journal of the Royal Statistical Society B*. 36:172-192

Ripley, B.D. (1976) Tests of "Randomness" for Spatial Point Patterns, *Journal of the Royal Statistical Society B*, 41:368-374

Ripley, B.D. (1981) *Spatial Statistics*. New York: Wiley.

Ripley, B.; Venables, W. (2015) Package 'class': Functions for Classification. R Package version 7.3-14 https://cran.r-project.org/web/packages/class/class.pdf

Ripley, B.D.; Venables, W. (2016) Package 'nnet' : Feed-Forward Neural Networks and Multinomial Log-Linear Models, R Package Version 7.3-12 https://cran.r-project.org/web/packages/nnet/nnet.pdf

Roche, H.; Brugal, J.P.; Delagnes, A.; Feibel, C.; Harmand, S.; Kibunjia, M.; Prat, S.; Texier, P.J. (2003) Les Sites Archéologiques Plio-Pléistocènes de la Formation de Nachukui, Ouest-Turkana, Kenya: Bilan Synthétique 1997-2001, *Comptes Rendus Palevol*. 2:663-673

Rodríguez-Álvarez, X.P. (1997) *Sistemas Técnicos de Producción Lítica del Pleistoceno Inferior y Medio de la Península Ibérica: Variabilidad Tecnológica entre Yacimientos del Noreste de la Sierra de Atapuerca*. PhD Thesis. Tarragona: Universidad Rovira i Virgili

Rodríguez-Hidalgo, A.; Saladié, P.; Marín, J.; Carbonell, E.; Canals, A. (2011) *Some Curious Spanish Tapas*. Salou: Hominid Carnivore Interactions.

Rodríguez-Hidalgo, A. (2015) *Dinámicas Subsistenciales durante el Pleistoceno Medio en la Sierra de Atapuerca: los Conjuntos Arqueológicos de TD10.1 y TD10.2*. PhD Thesis. Tarragona: Universidad Rovira i Virgili.

Rodríguez-Hidalgo, A.; Saladié, P.; Ollé, A.; Arsuaga, J.L.; Bermúdez de Castro, J.M.; Carbonell, E. (2017) Human Predatory Behavior and the Social Implications of Communal Hunting Based on Evidence from the TD10.2 Bison Bone Bed at Gran Dolina (Atapuerca, Spain), *Journal of Human Evolution*. 105:89-122

Rodríguez-Hidalgo, Antonio; García-Argudo, Gala; Morales, Juan I; Cebrià, Artur; Courtenay, Lloyd A.; Fernández-Marchena, Juan L.; Marín, Juan; Saladié, Palmira; Soto, María: Tejero, José-Miguel; Fullola, Josep María. A Châtelperronian Cut-Marked Raptor Phalange from Cova Foradada (Calfell, Spain). NeanderART 2018, *International Conference under the Aegis of UISPP and the auspices of IFRAO*. 22-26 August. Torino (Italy): Torino University, Campus Luigi Einaudi

Rodríguez-Hidalgo, Antonio; Morales, Juan I; Cebrià, A.; Courtenay, Lloyd A.; Fernández-Marchena, Juan L.; García-Argudo, Gala; Marín, Juan; Saladié, Palmira; Soto, María: Tejero, José-Miguel; Fullola, Josep María (Pre-Print Pending Publication) The Châtelperronian Neandertals of Cova Foradada (Calfell, Spain) used Iberian Imperial Eagle Phalanges for Symbolic Purposes, *PeerJ Preprints*. DOI: 10.7287/peerj.preprints.27133v1

Roe, D.A. (1981) *The Lower and Middle Palaeolithic Periods in Britain*. London: Routledge.

Rohlf, F.J.; Slice, D.E. (1990) Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks, *Systematic Zoology*. 39:40-59. DOI: 10.2307/2992207

Rohlf, F.J.; Marcus, L.F. (1993) A Revolution in Morphomerics, *Trends in Ecology & Evolution*. 8:129-132.

Rohlf, FK (1999) Shape Statistics: Procrustes Superimpositions and Tangent Spaces. *Journal of Classification*. 16(2):197-223.

Rohlf, F.K. (2000) Statistical Power Comparisons Among Alternative Morphometric Methods, *American Journal of Physical Anthropology*. 111:463-478

Rohlf, F.K. (2017) *tpsDig2 v.2.29*. New York: Ecology & Evolution and Anthropology, Stony Brook University. http://life.bio.sunysb.edu/morph/

Roopnarine, P.D. (2009) Ecological Modeling of Paleocommunity Food Webs, *Conservation Paleobiology*. 15:195-220.

Rosas, A.; Ríos, L.; Estalrrich, A.; Liversidge, H.; García-Tabernero, A.; Huguet, R.; Cardoso, H.; Bastir, M.; Lalueza-Fox, C.; Rasilla, M.; Dean, C. (2017) The Growth Pattern of Neandertals, Reconstruted from a Juvenile Skeleton from El Sidrón (Spain), *Science*. 357(6357):1282-1287. DOI: 10.1126/science.aan6463

Ross, A.H.; Slice, D.E; Williams, S.E. (2010) Geometric Morphometric Tools for the Classification of Human Skulls. Washington: *Department of Justice*. Document Number 231195. Award Number 2005-MU-BX-K078

Rougier, H; Crevecoeur, I; Beauval, C; Posth, C; Flas, D; Wising, C; Furtwängler, A; Germonpré, M; Gómez-Olivencia, A; Semal, P; Van der Plicht, J; Bocherens, H; Krause, J (2016) Neandertal Cannibalism and Neandertal Bones used as Tools in Nourthern Europe, *Scientific Reports*. 6:29005. DOI: 10.1038/srep29005

RStudio Team (2015) *RStudio: Integrated Development for R*. https://www.rstudio.com/.

Ruberto, C.D.; Putzu, L.; Rodriguez, G. (2018) Fast and Accurate Computation of Orthogonal Moments for Texture Analysis, *Pattern Recognition*. 83:498-510.

Rubio-Jara, S.; Panera, J.; Santonja, M.; Pérez-González, A.; Yravedra, J.; Domínguez-Rodrigo, M.; Bello, P.; Rojas, R.; Mabulla, A.; Baqudano, E. (2017) Site Function and Lithic Technology in the Acheulean Technocomplex: A Case Study from Thiongo Korongo (TK), Bed II, Olduvai Gorge, Tanzania, *BOREAS*. 46(4):849-917. DOI: 10.1111/bor.12275

Rufa, A.; Blasco, R.; Rosell, J.; Vaquero, M. (2017) What is Going on at the Molí del Salt Site? A Zooarchaeological Approach to the Last Hunter Gatherers from South Catalonia, *Hisotircal Biology*. 30:786-806. DOI: 10.1080/08912963.2017.1315685

Ruiter, DJ; Berger, LR (2000) Leopards as Taphonomic Agentsin Dolomitic Caves – Implications for Bone Accumulations in the Hominid-Bearing Deposits of South Africa, *Journal of Archaeological Science*. 27(8):665-684.

Russakovsky, O.; Deng, J.; Huang, Z.; Berg, A.; Fei-Fei, L. (2013) Detecting Avocados to Zucchinis: What we have done, and what we are doing. *ICCV*. http://ai.stanford.edu/~olga/papers/iccv13-ILSVRCanalysis.pdf

Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Koshla, A.; Bernstein, M.; Berg, A.C.; Fei-Fei, L. (2015) ImageNet Large Scale Visual Recognition Challenge, *IJCV*. http://www.image-net.org/challenges/LSVRC/

Russell, S.J.; Norvig, P. (2003) *Artificial Intelligence: a Modern Approach*. New Jersey: Upper Saddle River Press.

Sahle, Y.; Zaatari, S.; White, T.D. (2017) Hominid Butchers and Biting Crocodiles in the African Plio-Pleistocene, *PNAS*. 114(50):13164-13169. DOI: 10.1073/pnas.1716317114

Sahnouni, M.; Parés, J.M.; Duval, M.; Cáceres, I.; Harichane, Z.; van der Made, J.; Pérez-González, A.; Abdessadok, S.; Kandi, N.; Derradji, A.; Medig, M.; Boulaghraif, K.; Semaw, S. (2019) 1.9-Million-Year and 2.4-Million-Year-Old Artifacts and Stone Tool-Cutmarked bones from Ain Boucherit, Algeria, *Science*. 362:1297-1301

Sala, N.; Arsuaga, J.L.; Haynes, G. (2013) Taphonomic Comparison of Bone Modifications Caused by Wild and Captive Wolves (Canis lupus), *Quaternary International*. 330:126-135. DOI: 10.1016/j.quaint.2013.08.017

Salapaka, S.; Salapaka, M. (2008) Scanning Probe Microscopy, *IEEE Control Systems Magazine*. 28(2):65-83

Salimans, T:; Ho, J.; Chen, X.; Sidor, S.; Sutskever, I. (2017) Evolution Strategies as a Scalable Alternative to Reinforcement Learning, *Statistical Machine Learning*. arXiv: 1703.03864v2

Samuel, A.L. (1959) Some Studies in Machine Learning using the Game of Checkers, *IBM Journal of Research and Development*. 3(3):210-229

Sánchez-Yustos, P.; Diez-Martín, F.; Domínguez-Rodrigo, M.; Fraile, C.; Duque, J.; Uribelarrea, D.; Mabulla, A.; Baquedano, E. (2016) Techno-Economic Human Behavior in a Context of Recurrent Megafaunal Exploitation at 1.3 Ma, Evidence from BK4b (Upper Bed II, Olduvai Gorge, Tanzania), *Journal of Archaeological Science: Reports*. 9:386-404

Sánchez-Yustos, P.; Diez-Martín, F.; Domínguez-Rodrigo, M.; Duque, J.; Fraile, C.; Baquedano, E.; Mabulla, A. (2017a) Diversity and Significance of Core Preperation in the Developed Oldowan Technology: Reconstructing the Flaking Processes at SHK and BK (Middle Upper Bed II, Olduvai Gorge, Tanzania, *BOREAS*. 46(4):874-893. DOI: 10.1111/bor.12237

Sánchez-Yustos, P.; Diez-Martín, F.; Domínguez-Rodrigo, M.; Duque, J.; Fraile, C.; Díaz, I.; Francisco, S.; Baquedano, E.; Mabulla, A. (2017b) The Origin of the Acheulean. Technofunctional Study of the FLK-West Lithic Record (Olduvai, Tanzania), *PLoS ONE*. 12(8):e0179212. DOI: 10.1371/journal.pone.0179212

Sanchez-Yustos, P.; Diez-Martín, F.; Domínguez-Rodrigo, M.; Fraile, C.; Duque, J.; Díaz, I.; Francisco, S.; Baquedano, E.; Audax, M. (2018) Acheulean without Handaxes? Assemblage Variability at FLK West (Lowermost Bed II, Olduvai, Tanzania), *Journal of Anthropological Sciences Reports*. 96:53-73

Santonja, M.; Panera, J.; Rubio-Jara, S.; Pérez-González, A.; Uribelarrea, D.; Domínguez-Rodrigo, M.; Mabulla, A.Z.P.; Bunn, H.T.; Baquedano, E. (2014) Technological Strategies and the Economy of Raw Materials in the TK (Thiongo Korongo) Lower Occupation, Bed II, Olduvai Gorge, Tanzania, *Quaternary International*. 322-323:181-208

Santonja, M.; Rubio-Jara, S.; Panera, J.; Pérez-González, A.; Rojas-Mendoza, R.; Domínguez-Rodrigo, M.; Mabulla, A.Z.P.; Baquedano, E. (2018) Bifacial Shaping at the TK Aheulean Site (Bed II, Olduvai Gorge, Tanzania): New Excavations 50 Years after Mary Leakey. In: R. Gallotti, and M. Mussi (Eds.) *The Emergence of the Acheulean in East Africa and Beyond*. The Netherlands: Springer

Satopa, V.; Albrecht, J.; Irwin, D.; Raghavan, B. (2017) Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior, *Python PYPI*. https://github.com/arkevi/kneed.git

Schaller, G.B.; Lowther, G.R. (1969) The Relevance of Carnivore Behavior to the Study of Early Hominids. *SW Journal of Anthropology*. 25:307-341

Schuriman,, D.J. (1987) A Comparison of the Two One Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Biovariability, *Journal of Pharmacokinetic Biopharmacy*. 15:657-680

Seidler, H.; Falk, D.; Stringer, C.; Wilfing, H.; Müller, G.B.; Nedden, D.; Weber, G.W.; Reicheis, W.; Arsuage, J.L. (1997) A Comparative Study of Stereolithiographically Modelled Skulls of Petralona and Broken Hill: Implications for Future Studies of Middle Pleistocene Hominid Evolution, *Journal of Human Evolution*. 33:691-703 DOI: 10.1006/jhev.1997.0163

Selvaggio, M.M. (1994a) *Identifying the Timing and Sequence of Hominid and Carnivore Involvement with Plio-Pleistocene Bone Assemblages from Carnivore Tooth Marks and Stone Tool Butchery Marks on Bone Surfaces*. PhD Thesis. New Brunswick: Rutgers University.

Selvaggio, M.M. (1994b) Carnivore Tooth Marks and Stone Tool Butchery Marks on Scavenging Bones: Archaeological Implications, *Journal of Human Evolution*. 27:215-228

Selvaggio, M.M. (1998) The Archaeological Implications of Water Catched Hyena Kills, *Current Anthropology*. 39(3):380-383. DOI: 10.1086/204750

Selvaggio, M.M.; Wilder, J. (2001) Identifying the Involvement of Multiple Carnivore Taxa with Archaeological Bone Assemblages, *Journal of Archaeological Science*. 28(5):465-470

Semaw, S.; Roberts, M.J.; Quade, J.; Renne, P.R.; Butler, R.F.; DOmínguez-Rodrigo, M.; Stout, D.; Hart, W.S.; Pickering, T.; Simpson, S.W. (2003) 2.6 Million Year Old Stone Tools and Associated Bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia, *Journal of Human Evolution*. 45:169-177

Semaw, S., Rogers, M.J.; Stout, D. (2009) The Oldowan-Acheulian transition: is there a "Developeded Oldowan" artifact tradition? In: M. Camps and P. Chauzan (Eds.):

Sourcebook of Paleolithic Transitions Methods, Theories, and Interpretations. New York: Springer. 173-193

Semenov, S.A. (1964) *Prehistoric Technology. An Experimental Study of the Oldest Tools and Artifacts from Traces of Manufacture and Wear*. London: Cory, Adams and Mackay Ltd.

Schick, K.; Toth, N. (1998) *Making Silent Stones Speak: Human Evolution and the Dawn of Technology*. New York: Simon and Schuster.

Schick, K.; Toth, N. (2006) *The Oldowan: Case Studies into the Earliest Stone Age.* Gosport: Stone Age Institute Press

Schiffer, M.B. (1987) *Formation Processes of the Archaeological Record.* New Mexico: University of New Mexico Press.

Shanks, M.; Tilley, C. (1987a) *Social Theory and Archaeology*. Cambridge: Polity Press.

Shanks, M.; Tilley, C. (1987b) *Re-Constructing Archaeology: Theory and Practice*. Cambridge

Sherafatian, M. (2018) Tree-Based Machine Learning Algorithms Identified Minimal Set of miRNA Biomarkers for Breast Cancer Diagnosis and Molecular Subtyping, *Gene*. 677:111-118

Shipman, P.; Rose, J. (1983) Evidence of Butchery and Hominid Activities at Torralba and Ambrona; An Evaluation Using Microscopic Techniques. *Journal of Archaeological Science*. 10(5):465-474.

Shipman, P.; Fisher, D.C.; Rose, J.J. (1984a) Mastodon butchery: microscopic evidence of carcass processing and bone tool use. *Paleobiology* 10 (3):358–365.

Shipman, P.; Foster, G.; Schoeninger, M. (1984b) Burnt bones and teeth: an experimental study of color, morphology, crystal structure and shrinkage. *Journal of Archaeological Science* 11:307–325.

Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; Lillicrap, T.; Simonyan, K.; Hassabis, D. (2017) Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. *Algortihm*. arXiv: 1712.01815v1

Simonite, T. (2018) The DIY Tinkers Harnessing the Power of Artificial Intelligence. *WIRED* 26(5).

Simonyan, K.: Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition, *ICLR*. arXiv: 1409.1556v6

Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T. (2005) ROCR: Visualizing Classifier Performance in R, *Bioinformatics*. 21(20):3940-3941

Slice, DE (2001) Landmark Coordinates Aligned by Procrustes Analysis do not Lie in Kendall's Shape Space. *Systematic Biology*. 50(1):141-149.

Solem, J.E. (2012) *Programming Computer Vision with Python: Tools and Algorithms for Analyzing Images*. Tokyo: O'Reilly

Spennemann, D.H.R. (1990) Don't forget the bamboo. On recognising and interpreting butchery marks in tropical faunal assemblages, some comments asking for caution, problem solving in taphonomy. *Archaeological and palaeontological studies from Europe, Africa and Oceania* 2:80–101

Speth, J.D. (1989) Early Hominid Hunting and Scavenging: the Role of Meat as an Energy Source. *Journal of Human Evolution*. 18:329-343

Speth, J.D. (2013) *Bison Kills and Bone Counts*. Chicago: University of Chicago Press.

Spizhevoy, A. (2018) *OpenCV3 Computer Vision with Python Cookbook*. Mumbai: Packt

Stanford, C.B.; Bunn, H.T. (2001) *Meat Eating and Human Evolution*. Oxford: Oxford University Press.

Stapert, D.; Street, M. (1997) High Resolution or Optimum Resolution? Spatial Analysis of the *Federmesser* site at Andernach, Germany. *World Archaeology*. 29(2):172-194

Steele, D.G.; Baker, B.W. (1993) Multiple Predation: a Definitive Human Hunting Strategy. In: J. Hudson (Ed.) *From Bones to Behavior*. *Ethnoarchaeological and Experimental Contributions to the Interpretation of Faunal Remains*. Carbondale: Center for Archaeological Investigation. 9-37

Steiger, R.H.; Jäger, E. (1977) Subcomission on Geochronology: Convention on the Use of Decay Constants in Geo and Cosmochronology. *Earth and Planetary Science Letters.* 36:359-362.

Stemp, W.J.; Watson, A.S.; Evans, A.A. (2015) Surface Analysis of Stone and Bone Tools. *Surface Topography: Metrology and Properties*. DOI: 10.1088/2051-672X/4/1/013001

Stiner, MC (2002) On in situ Attrition and Vertebrate Body Part Profiles, *Journal of Archaeological Science*. 29(9):979-991

Stinnesbeck, S.R., Stinnesbeck, W., Terrzas Mata, A., Avilés Olguín, J., Benavente Sanvicente, M., Zell, P., Frey, E., Lindauer, S., Sandoval, C.R., Velázquez Morlet, A., Acevez Nuñez, E., González González, A., 2018. The Muknal Cave near Talum, Mexico: an Early-Holocene Funeral Site on the Yucután Peninsula. *The Holocene.* DOI: 10.1177/0959683618798124

Stout D.; Toth N.; Shick K.; Stout J.; Hutchins G. (2000) Stone Tool-Making and Brain Activation: Position Emission Tomography (PET) Studies, *Journal of Archaeological Science*. 27:1215-1223

Stout, D.; Chaminade, T. (2007) The Evolutionary Neuroscience of Tool Making, *Neuropsychologia*. 45:1091-1100

Stout D.; Toth N.; Schick K.; Chaminade T. (2008) Neural Correlates of Early Stone Age Toolmaking: Technology, Language and Cognition in Human Evolution, *Philosophical Transactions of the Royal Society B.* 363(1499):1939-1949. DOI: 10.1098/rstb.2008.0001

Stout, D.; Chaminade, T. (2012) Stone Tools, Language and the Brain in Human Evolution, *Phil. Trans. R. Soc. B.* 367:75-87. DOI: 10.1098/rstb.2011.0099

Stout, D.; Hecht, E.; Khrelsheh, N.; Bradley, B.; Chaminade, T. (2015) Cognitive Demands of Lower Palaeolithic Toolmaking, *PLoS ONE*. 10(4):e0128256. DOI: 10.1371/journal.pone.0121804

Stringer, CB (1985) The Hominid Remains from Gough's Cave, *Proceedings of the University of Bristol Spelaeological Society*. 17(2):145-152.

Stünkel, I. (2017) Hippopotami in Ancient Egypt, *Heilbrunn Timeline of Art History.* New York: The Metropolitan Museum of Art. [Online] Available from: www.metmuseum.org/toah/hd/hipi/hd_hipi.htm [Accessed: 30/04/2019]

Su, H.; Deng, J.; Fei-Fei, L. (2012) Crowdsourcing Annotations for Visual Object Detection, *Human Computation Workshop*. http://ai.stanford.edu/~haosu/bbox_submission.pdf

Sutcliffe, A.J. (1970) Spotted Hyaena: Crusher, Gnawer, Digester and Colector of Bones. *Nature* 227:1110-1113

Suwa, G.; Nakaya, H.; Asfaw, B.; Saegusa, H.; Amzaye, A.; Kono, R.T.; Beyene, Y.; Katoh, S. (2003) Plio-Pleistocene Terrestrial Mamal Assemblage from Konso, Southern Ethiopia, *Journal of Vertebrate Paleontology*. 23(4):901-916

Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. (2015) Going Deeper with Convolutions, *Computer Vision Foundation*. arXiv: 1409.4842

Szepesvári, C. (2009) Algorithms for Reinforcement Learning, *Synthesis Lectures on Artificial Intelligence an dMachine Learning.*

Tafti, A.P.; Kirkpatrick, A.B.; Alavi, Z.; Owen, H.A.; Yu, Z. (2015) Recent Advances in SEM Surface Reconstruction, *Micron*. 78:54-66. DOI: 10.1016/j.micro.2015.07.005

Tapak, L.; Shirmohammadi-Khorram, N.; Amini, P.; Alafchi, B.; Hamidi, O.; Poorolajal, J. (2018) Prediction of Survival and Metastasis in Breast Cancer Patients using Machine Learning Classifiers, *Clinical Epidemiology and Global Health*. DOI: 10.1016/j.cegh.2018.10.003

Tegmark, M. (2017) *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York: Borzoi Books.

Theano Development Team (2017) *Theano: A Python Framework for Fast Computation of Mathematical Expresssions*. https://deeplearning.net/software/theano

Thompson, C. (2018) How to Teach Aritifical Intelligence some Common Sense, *Wired*. 12:74-89

Thompson, D.A.W. (1915) Mophology and Mathematics, *Transactions of the Royal Society of Edinburgh*, 50:857-895

Thompson, D.A.W. (1917) *On Growth an-d Form*. Cambridge: Cambridge University Press.

Thompson, D.A.W. (1981) *Miseria de la Filosofia*. Barcelona: Crítica.

Thompson, J.C.; McPherron, S.P.; Bobe, R.; Reed, D.; Barr, W.A.; Wynn, J.G.; Marean, C.W.; Geraads, D.; Alemseged, Z. (2015) Taphonomy of Fossils from the Hominin-Bearing Deposits at Dikika, Ethiopia, *Journal of Human Evolution*. 86:112-135 DOI: 10.1016/j.jhevol.2015.06.013

Thompson, J.C.; Carvalho, S.; Marean, C.W.; Alemseged, Z. (2019) Origins of the Human Predatory Pattern: The Transition to Large Animal Exploitation by Early Hominins, *Current Anthropology*. 60(1):1-23

Throndike, R.L. (1953) Who Belongs in the Family?, *Psychometrika*. 18(4):267-276

Tobias, P.V. (1967) *Olduvai Gorge. Vol.2. The Cranium and Maxillary Dentition of Australopithecus (Zinjanthropus) Boisei*. New York: Cambridge University Press.

Tobias, P.V. (1991) *Olduvai Gorge: The Skulls, Endocasts and Teeth of Homo habilis*. Cambridge: Cambridge University Press

Tomasello, M.; Davis-Dasilva, M.; Camak, L.; Bard, K. (1987) Observational Learning of Tool Use by Young Chimpanzees, *Human Evolution*. 2(2):175-183.

de la Torre, I.; Mora, R. (2013) The transition to the Acheulean in East Africa: an assessment of paradigms and evidence from Olduvai Gorge (Tanzania). *Journal of Archaeological Method and Theory*. 21(4):781-823. DOI: 10.1007/s10816-013-9176-5

de la Torre, I.; Benito-Calvo, A.; Arroyo, A.; Zupancich, A.; Proffitt, A. (2013) Experimental protocols for the study of battered stone anvils from Olduvai Gorge (Tanzania). *Journal of Archaeological Science*. 40:313–332

de la Torre, I. ;Mora, R. (2014) The transition to the Acheulean in East Africa: an assessment of paradigms and evidence from Olduvai Gorge (Tanzania). *Journal of Archaeological Method and Theory*. 21:781–823.

de la Torre, I.; Mora, I. (2018) Oldowan Technological Behaviour at HWK EE (Olduvai Gorge, Tanzania), *Journal of Human Evolution*. 120:236-273

Tosun, A.B.; Kandemir, M.; Sokmensuer, C.; Gunduz-Demir, C. (2009) Object-Oriented Texture Analysis for the Unsupervised Segmentation of Biopsy Images for Cancer Detection, *Pattern Recognition*. 42:1104-1112

Toth, N (1985) The Oldowan Reassessed: a Close Look at Early Stone Artifacts, *Journal of Archaeological Science*. 12:101-120

Toth, N.; Schick, K. (2009a) The Importance of Actualistic Studies in Early Stone Age Research: Some Personal Reflections. In: K. Schick and N. Toth (Eds.) *The Cutting Edge: New Approaches to the Archaeology of Human Origins*. Gosport: Stone Age Institute Press. 267-344

Toth, N.; Schick, K. (2009b) The Oldowan Tool Making of Early Hominins and Chimpanzees Compared, *Annual Review of Anthropolgy*. 38:289-305

Toth N.; Schick K. (2018) An Overview of the Cognitive Implications of the Oldowan Industrial Complex, *Azania: Archaeological Research in Africa.* 53(1):3-39

Turner, A. (1992) Large Carnivores and Earliest European Hominids: Changing Determinants of Resource Availability during the Lower and Middle Pleistocene, *Journal of Human Evolution*. 22:109-126

Turner, R. (2013) *Flint Knapping: A Guide to Making your own Stone Age Toolkit.* Gloucestershire: The History Press.

Ullrich, H (1989) Neanderthal Remains from Krapina and Vindija Mortuary Practices, Burials or Cannibalism? *Humanbiologie Budapest*. 19:15-19.

Ullrich, H (2005) Cannibalistic Rites Within Mortuary Practices from the Paleolithic to Middle Ages in Europe, *Anthropologie*. 43(2-3):249-261.

Underdown, S.J.; Kumar, K.; Houldcroft, C. (2017) Network Analysis of the Hominin Origin of Herpes Simplex virus 2 from Fossil Data, *Virus Evolution*. 3(2):vex026 DOI: 10.1093/ve/vex026

Ungar, P.S.; Grine, F.E.; Teaford, M.F. (2006) Diet in Early *Homo:* a Review of the Evidence and a New Model of Adaptive Versatility, *Annual Review of Anthropology*. 35:209-228

Uribelarrea, D.; Domínguez-Rodrigo, M.; Pérez-González, A.; Vegas Salamanca, J.; Baquedano, E.; Mabulla, A.; Musiba, C.; Barboni, D.; Cobo-Sánchez, L. (2014) Geo-Archaeological and Geometrically Corrected Reconstruction of the 1.84 Ma FLK-Zinj Paleolandscape at Olduvai Gorge, Tanzania, *Quaternary International*. 322-323:7-31

Uribelarrea, D.; Domínguez-Rodrigo, M. (2017) Geochronology in a Meandering River: a Study of the BK Site (1.35 Ma) Upper Bed II, Olduvai Gorge (Tanzania), *Palaeogeography, Palaeoclimatology, Palaeoecology*. 488:76-83

Uribelarrea, D.; Martín-Perea, D.; Díez-Martín, F.; Sánchez-Yustaos, P.; Domínguez-Rodrigo, M.; Baquedano, E:; Mabulla, A. (2017) REconstruction of the Plaeolandscape during the Earliest Acheulian of FLK West: The Co-Existence of Oldowan and Acheulian Industries during Lowermost Bed II (Olduvai Gorge, Tanzania), *Palaeogeography, Palaeoclimatology, Palaeoecology*. 488:50-58. DOI: 10.1016/j.palaeo.2017.04.014

Uribellarea, D. (2018) La Geología de FLK-West. In: F. Diéz-Martín (Ed.) *En África Hace 1.7 Millones de Años: El Origen del Achelense*. Madrid: Museo Arqueológico Regional. 41-52

Uerpmann, H.P. (1973) Animal Bone Finds and Economic Archaeology: A Critical Study of "Osteo-Archaeological" Method, *World Archaeology*. 4(3):307-322.

Valensi, P; Crégut-Bonnoure, E; Defleur, A (2012) Archaeozoological Data from the Mousterian Level from Moula-Guercy (Ardèche, France) Bearing Cannibalised Neanderthal Remains, *Quaternary International*. 252:48-55.

Van Gelder L (2010) 10 years in Rouffignac: a collective report on findings from a decade of finger flutings research, *IFRAO congress, September 2010 – Symposium: Pleistocene Art in Europe (Pre-Acts)*.

Van Gelder L (2014) Paleolithic finger flutings and the question of writing. *Time and Mind* 7(2):141–153

Van Gelder L (2015) Counting the children: the role of children in the production of finger flutings in four upper Palaeolithic caves. *Oxford Journal of Archaeology*. 34(2):117–138

Van Gelder L, Sharpe K (2015) Evidence for cave marking by Palaeolithic children. *Antiquity*. 80(310):937–947

Venables, W.N.; Ripley, B.D. (2002) *Modern Applied Statstics.* New York: Springer

Vergès. J.M. (2003) *Caracterització dels models d'instrumental lític del mode 1 a partir de les dades de l'analisi funcional dels conjunts litotènics d'Aïn Hanech i El-Kherba (Algèria), Monte Poggiolo i Isernia la Pineta (Itàlia).* PhD Thesis. Tarragona: Universitat Rovira i Virgili.

Vergès. J.M.; Morales, J.I. (2014) The Gigapixel Image Concept for Graphic SEM Documentation. Applications in Archaeological Use-Wear Studies, *Micron* 65:15-19. DOI: 10.1016/j.micron.2014.04.009

Villa, P.; Mahieu, E: (1991) Breakage Patterns of Human Long Bones, *Journal of Human Evolution*. 21:27-48. DOI: 10.1016/0047-2484(91)90034-S

Vinuesa, V.; Iurino, D.A.; Madurell-Malapeira, J.; Liu, J.; Fortuny, J.; Sardella, R.; Alba, D.M. (2015) Inferences of Social Behavior in Bone-Cracking Hyaenids (Carnivore, Hyaenidae) Based on Digital Paleoneurological Techniques: Implications for Human-Carnivoran Interactions in the Pleistocene, *Quaternary International*. 413(B):7-14 DOI: 10.1016/j.quaint.2015.10.037

Visualization Sciences Group (2019) *Amira 5.0.* USA: Visualization Sciences Group. https://www.visageimaging.com

Walden, S.J.; Evans, S.L.; Mulville, J. (2017) Changes in Vickers Hardness During the Decomposition of Bone: Possibilities for Forensic Anthropology, *Journal of Mechanical Behaviour Biomedical Materials.* 65:672-678

Walker, P.L.; Long, J.C. (1977) An Experimental Study of the Morphological Characteristics of Tool Marks, *American Antiquity*. 42(4):605-616

Wallduck, R.; Bello, S.M. (2018) Cut Mark Micro-Morphometrics Associated with the Stage of Carcass Decay: a Pilot Study using Three-Dimensional Microscopy, *Journal of Archaeological Science Reports*.18:174-185

van der Walt, S.; Colbert, C.: Varpquaux, G. (2011) The NumPy Array: a Structure for Efficient Numerical Computation, *Computing in Science & Engineering.* 13:22-30 DOI: 10.1109/MCSE.2011.37

Walter, M.J.; Trauth, M.H (2013) A Matlab Based Operation Analysis of Acheulean Handaxe Accumulations in Olorgesailie and Kariandusi, Kenya Rift, *Journal of Human Evolution*. 64(6):569-581. DOI: 10.1016/j.jhevol.2013.02.011

Walter, R.C.; Manega, P.C.; Hay, R.L.; Drake, R.E.; Curtis, G.H. (1991) Laser-Fusion 40Ar/39Ar Dating of Bed I Olduvai Gorge, Tanzania, *Nature*. 354:145-149

Wang, S.; Chaovalitwongse, W.; Bubuška, R. (2012) Machine Learning Algorithms in Bipedal Robot Control, *Transactions on Systems, Man and Cybernetics*. 42(5):728-743

Washburn, S.L. (1963) Behavior and Human Evolution. In S.L. Washburn (Ed.) *Classification and Human Evolution.* Chicago: Aldine. 190-203

Wauthoz, B.; Dorning, K.J.; Hérissé, A.L.E. (2003) Crassiangulina variacornuta sp. nov. from the Late Llandovery and its Bearing on Silurian and Devonian Acritarch Taxonomy, *Bulletin de la Société Géologique de France*. 174(1):67-81.

Webb, G.I.; Agar, J.W.M. (1992) Inducing Diagnostic Rules for Glomerular Disease with the DLG Machine Learning Algorithm, *Artificial Intelligence in Medicine*. 4:419-430

West, J.; Louys, J (2007) Differentiating Bamboo from Stone Tool Cut Marks in the Zooarchaeological Record, with a Discussion on the use of Bamboo Knives, *Journal of Archaeological Science*. 34:512-518

Westaway, M.C.; Thompson, J.C.; Wood, W.B.; Njau, J. (2011) Crocodile Ecology and Taphonomy of Early Australasian Sites, *Environmental Archaeology*. 16(2):124-136. DOI: 10.1179/174963111X13110803260930

Weston, E.; Szabó, K.; Stern, N (2015) Pleistocene Shell Tools from Lake Mungo Lunette, Australia: Identification and Interpretation Drawing on Experimental Archaeology, *Quaternary International*. 427(A):229-242. DOI: 10.1016/j.quaint.2015.11.048

Whewell, W. (1847) *Of The Fundamental Antithesis of Philosophy, The Philosophy of the Inductive Sciences Founded Upon their History*. London: John W Parker. 2:16-50.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wierer, U.; Arrighi, S.; Bertola, S.; Kaufmann, G.; Baumgarten, B.; Pedrotti, A.; Pernter, P.; Pelegrin, J. (2018) The Iceman's Lithic Toolkit: Raw Material, Technology, Typology and Use. *PLoS ONE* 13(6):e0198292. DOI: 10.1371/journal.pone.0198292

Wijffels, J. (2017) *Computer Vision Algorithms for R Users*. Brussels: BNOSAC

Wilson, M.V.G. (1988) Taphonomic Processes: Information Loss and Information Gain. *Geoscience Canada*. 15:131-148.

Wilson, D.E.; Reeder, D.M. (2005) *Mammal Species of the World. A Taxonomic and Geographical*. Baltimore: Hopkins University Press. 3rd Edition.

White, J.R. (1979) Sequencing in-site Taphonomic Processes: The Lesson of the Eaton Briquets, *Midcontinental Journal of Archaeology*. 4(2):209-220.

Wolpoff, MH; Caspari, R (2006) Dies Krapina Reflect Early Neanderthal Paleodemography?, *Periodicum Biologorum*. 108:425-432.

Wood, B.; Richmond, B. (2002) Human Evolution: Taxonomy and Paleobiology, *Journal of Anatomy*. 197(1). DOI: 10.1046/j.1469-7580.2000.19710019.x

Wynn, T.; Tierson, F. (1990) Regional Comparison of the Shapes of Later Acheulian Handaxes, *American Anthropologist*. 92:73-87

Xhauflair, H.; Pawlik, A.; Forestier, H.; Saos, T.; Dizon, E.; Gaillard, C. (2017) Use-Related or Contamination? Residue and Use-Wear Mapping on Stone Tools used or Experimental Processing of Plants from Southeast Asia, *Quaternary International.* 427:80-93

Yalden, D.W.; Largen, M.J.; Kock, D. (1986) Catalogue of the Mammals of Ethiopia, *Monitore Zoologico Italiano*. 21(1):31-103. DOI: 10.1080/03749444.1986.10736707

Yang, Y.; Ye, Z.; Su, Y.; Zhao, Q.; Li, X.; Ouyang, D. (2018) Deep Learning for *in vitro* Predition of Pharmaceutical Formulations, *Acta Pharmaceutica Sinica B*. 9(1):117-185. DOI: 10.1016/j.apsb.2018.09.010

Yravedra, J. (2005) *Patrones de Aprovechamiento de Recursos Animales en el Pleistoceno Superior de la Península Ibérica: Estudio Tafonómico y Zooarqueológico de los Yacimientos del Esquilleu, Amalda, Cueva Ambrosio y la Peña de Estebanvela*. PhD Thesis. Madrid: Universidad Nacional de Educación a Distancia (UNED)

Yravedra, J. (2006) *Tafonomía Aplicada a Zooarqueología*. Madrid: Lerko Print.

Yravedra, J.; Lagos, L.; Bárcena, F. (2011) A Taphonomic Study of Wild Wolf (Canis lupus) Modifiaction of Horse Bones in Northwestern Spain, Journal of Taphonomy. 9(1):37-65.

Yravedra, J.; Lagos, L.; Bárcena, F. (2012) The Wild Wolf (Canis lupus) as a Dispersal Agent of Animal Carcasses in Northwestern Spain, *Journal of Taphonomy*. 10:291-238

Yravedra, J.; Andrés, M.; Domínguez-Rodrigo, M. (2013) A Taphonomic Study of the African Wild Dog (*Lycaon pictus*), *Journal of Archaeological Anthropological Science* 6:113-124. DOI: 10.1007/s12520-013-0164-1

Yravedra, J.; Diez-Martín, F.; Egeland, C.P.; Maté-González, M.Á.; Palomeque-González, J.F.; Arriaza, M.C.; Aramendi, J.; García Vargas, E.; Estaca-Gómez, V.; Sánchez, P.; Fraile, C.; Duque, J.; de Francisco Rodríguez, S.; González-Aguilera, D.; Uribelarrea, D.;

Mabulla, A.; Baquedano E.; Domínguez-Rodrigo, M. (2017a) FLK-West (Lower Bed II, Olduvai Gorge, Tanzania): a newearly Acheulean site with evidence for human exploitation of fauna. *Boreas*. 46(3):486-502. DOI: 10.1111/bor.12243

Yravedra, J.; Maté-González, M.A.; Palomeque-González, J.F.; Aramendi, J.; Estaca-Gómez, V.; Blazquez, M.S.; García Vargas, E.; Organista, E.; González-Aguilera, D.; Arriaza, M.C.; Cobo-Sánchez, L.; Gidna, A.; Uribelarrea del Val, D.; Baquedano, E.; Mabulla, A.; Domínguez-Rodrigo, M. (2017b) A New Approach to Raw Material use in the Exploitation of Animal Carcasses at BK (Upper Bed II, Olduvai Gorge, Tanzania): a Micro-Photogrammetric and Geometric Morphometric Analysis of Fossil Cut Marks, *BOREAS*. 46(4):860-873. DOI: 10.1111/bor.12224

Yravedra, J.; García Vargas, E.; Maté González, M.A.; Aramendi, J.; Palomeque-González, J.; Vallés-Iriso, J.; Matasanz-Vicente, J.; González-Aguilera, D.; Domínguez-Rodrigo, M. (2017c) The use of Micro-Photogrammetry and Geometric Morphometrics for identifying carnivore agency in bone assemblage. *Journal of Archaeological Science Reports* 14:106–115

Yravedra, J.; Domínguez-Rodrigo, M. (2018) Zooarqueología y Tafonomía en FLK-West. In: F. Diez-Martín (Ed.) *En África Hace 1,7 Millones de Años: El Origen del Achelense*. Madrid: Museo Arqueológico Regional. 72-88

Yravedra, J.; Aramendi, J.; Maté-González, M.Á.; Courtenay, L.A.; González-Aguilera, D. (2018) Differentiating Percussion Pits and Carnivore Tooth Pits using 3D Reconstructions and Geometric Morphometrics, *PLoS ONE*. 13(3):e0194324. DOI: 10.1371/journal.pone.0194324

Yravedra, J.; Maté-González, M.Á.; Courtenay, L.A.; López-Cisneros, P.; Estaca-Gómez, V.; Aramendia, J.; Andrés-Herrero, M.; Linares-Matás, G.; González-Aguilera, D.; Álvarez-Alonso, D. (2019) Approaching Raw Material Functionality in the Upper Magdalenian of Coímbre Cave (Asturias, Spain) through Geometric Morphometrics. *Quaternary International.* DOI: 10.1016/j.quaint.2019.01.008

Yustos, M; Yravedra, J (2015) Cannibalism in the Neanderthal World: an Exhaustive Revision, *Journal of Taphonomy*. 13(1):33-52.

Yu-Wei, C. (2015) *Machine Learning with R Cookbook*. Birmingham: Packt Publishing

Zhang, Z.; Liu, S.; Mei, X.; Xiao, B.; Zheng, L. (2017) Learning Completed Discriminative Local Features for Texture Classification, *Pattern Recognition*. 67:263-275

Zhang, J.; Kalantidis, Y.; Rohrbach, M.; Paluri, M. (2019) Large-Scale Visual Relationship Understanding, *Conference on Artificial Intelligence*. [Online] Available from: https://research.fb.com/publications/large-scale-visual-relationship-understanding/ [Accessed 26/02/2019]

Zheng, A.; Casari, A. (2018) *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. Tokyo: O'Reilly

Zollikofer, C.P.E.; Ponce de León, M.S. (2004) Kinematics of Cranial Ontogeny: Heterotopy, Heterchrony and Geometric Morphometric Analysis of Growth Models, *Journal of Experimental Zoology*. 302(B):322-340

Zollikofer, C.P.E.; Ponce de León, M.S. (2010) The Evolution of Hominin Ontogenies, *Seminars in Cell and Developmental Biology*. 21:441-452

Appendices

Appendix 1 – Abbreviations and Definitions

Appendix 1a
Abbreviations for the Named Localities in the Olduvai Gorge

This list presents in alphabetical order, all the named localities mentioned within this Master's Thesis and their abbreviations. The localities found within the gorge are usually named after the person making the first discovery, followed by the letters C, for cliff, K for Korongo, G, for Gully or S, for site. These can then be followed by a letter indicating their geographical location (N, North, E, East, S, South or W, West) with respect to the main site of each Korongo. FLK-N is thus North of FLK. FLK-NN is North of the FLK-N site, and so on. Finally, for any abbreviation displayed as OH followed by a number, for example OH5, this refers to hominid fossil remains, as Olduvai Hominid.

It is important to point out, however, that not all the sites found within the Gorge have been included in this list, only those mentioned within this Master's Thesis. For a full list and description of the named localities, consult M. Leakey (1965).

AMK – Amin Mturi Korongo
BK – Bell's Korongo
CK – Camp Korongo
CMK – Catherine Martin Korongo
Croc. K – Crocodile Korongo
DC – Donald's Cliff
DGS – Dorothy Garrod Site
DS – David's Site
DK – Douglas (Leakey) Korongo
EF-HR – Evelyn Fuchs-Hans Reck
Elephant K – Elephant Korongo
FC – Fuch's Cliff
FLK – Frida Leakey Korongo
GC – Abbreviation not explained by M. Leakey (1965)
GRC – Abbreviation not explained by M. Leakey (1965)
HG – Hoopoe Gully
HEG – Herberer's Gully
HK – Hopwood's Korongo
HWK – Henrietta Wilfrida Korongo
JK – Juma's Korongo
KK – Kudu Korongo (Also the name of the fault found crossing the Korongo)
Long Korongo – Long Korongo
LLK – Louis Leakey Korongo
MK – MacInnes Korongo
MCK – Margaret Cropper Korongo
MLK – Mary Leakey Korongo
MNK – Mary Nicol Korongo
OCS – Olduvai Carnivore Site
PDK – Peter Davies Korongo
PLK – Philip Leakey Korongo
PTK – Philip Tobias Korongo
SHK – Sam Howard Korongo
SWK – Sam White Korongo
TK – Thiongo Korongo
VEK – Vivian Evelyn (Fuchs) Korongo
WK – Wayland's Korongo

Appendix 1b
General Abbreviations

This list presents all the general abbreviations used throughout this study. These are additionally grouped into categories concerning the topic they are associated to.

Taphonomy

ATI – Angle of Tool Impact
BSM – Bone Surface Modifications
D – Depth
LDC – Left Depth of Incision at Convergent
MNE – Minimum Number of Elements
MNI – Minimum Number of Individuals
NE – Number of Elements
NISP – Number of Identifiable Specimens
NR – Number of Remains
OA – Opening Angle
RDC – Right Depth of Incision at Convergent
WIB – Width of Incision at Base
WIM – Width of Incision at Middle
WIS – Width of Incision at Surface

Statistics and Geometric Morphometrics

AIC – Akaike Information Criterion
ANOVA – Analysis of Variance
CV Scores – Canonical Variance Scores
CVA – Canonical Variance Analysis
EM – Expectation Maximisation
GMM – Geometric Morphometrics
GPA – Generalised Procrustes Analysis
LDA – Linear Discriminant Analysis
MANOVA – Multivariate Analysis
MCA – Multiple Correspondence Analysis
PC Scores – Principal Components Scores
PCA – Principal Components Analysis
TOST – Two-One Sided Equivalence Test
TPS – Thin Plate Spline

Artificial Intelligence

AI – Artificial Intelligence
AUC – Area Under Curve
C5.0 – A decision tree algorithm
CNN – Convolutional Neural Network
CTREE – Conditional Inference Tree
CV – Computer Vision
DL – Deep Learning
DNN – Deep Neural Network
KNN – K-Nearest Neighbour
MDA – Mixed Discriminant Analysis
ML – Machine Learning
NB – Naïve Bayes

NNET – Neural Network
OOB – Out-of-Bag
PLSDA – Partial Least Square Discriminant Analysis
RBF – Radial Basis Function
RF – Random Forest
ROC – Receiver Operating Characteristic
SDR – Standard Deviation Reduction
SVM – Support Vector Machine

Miscellaneous Abbreviations

BCE – Before Common Era
BP – Before Present
CAI – Centro de Asistencia a la Investigación
CE – Common Era
CPU – Central Processing Unit
DO – Developed Oldowan
ESEM – Environmental Scanning Electron Microscopy
FOV – Field Of View
IPHES – Institut Catalá de Paleoecologia Humana I Evolució Social
LAS – Lower Augitic Sandstone
LCT – Large Cutting Tool
RAM – Random Access Memory
SEM – Scanning Electron Microscopy
TEM – Transmission Electron Microscopy
TOPPP – The Olduvai Palaeoanthropological and Palaeoecological Project
VS – Visual Studio

Appendix 1c
Glossary of Geometric Morphometric Terms

This list presents definitions to some of the key terms in geometric morphometrics. Many of these definitions have been adapted from Mitteroecker and Gunz (2009).

Bending energy	The amount of deformation between two landmark configurations, defined through deformation grids and thin plate spline funcitons.
Cartesian coordinates	Coordinates that indicate the location of a point relevant to a fixed point of origin.
Centroid size	The measurement of size in Procrustes analysis.
Deformation grids	A grid placed over the homologous points in landmark configuration models which deforms based on differences between each configuration.
Euclidian distance	Distance between points in a transformed feature space.
Form	The geometric properties of an object that are invariant to rotation and translation (shape plus overall size). Definition by Mitteroecker and Gunz (2009).
Generalised Procrustes Analysis	Superimposition of landmark configurations that can be used to compute shape coordinates.
Geometric homology	The assumption that landmarks are homologous points that can be located on all specimens.

Kendall shape space	A non-linear space that is induced by a series of shape coordinates.
Landmark	A homologous point of correspondence on each individual that matches between and within populations. Adapted from Dryden and Mardia (1998).
Mahalanobis distance	The distance between a series of distributions in a constructed feature space.
Procrustes distance	Euclidean Distance between configurations of landmark coordinates, providing a quantitative measure of differences between shapes.
Procrustes Shape Coordinates	Landmark coordinates obtained through Procrustes superimposition (see Generalised Procrustes Analysis) representing the shape of an individual
Semilandmark	Landmarks that are computed to fit along smooth surfaces and curves.
Shape	The geometric properties of an object that are invariant to scale, rotation and translation. Definition by Mitteroecker and Gunz (2009).
Tangent space	Euclidean approximation of Kendall shape space.
Thin Plate Spline	A function that models differences between shapes as a smooth deformation. This is used to compute deformation grids.

Appendix 1d
Glossary of Statistics and Artificial Intelligence Terms

This list presents definitions to some of the key terms in statistics and Artificial Intelligence.

Back propagation	A supervised learning method which uses the error produced in the output signal to adjusting the weights accordingly and improve the consequent results during training.
Batch size	The number of training examples used in one iteration (see epoch)
Bayesian analysis	An approach to statistically calculating probabilities of an event based on new data obtained about this event.
Bootstrap	A random resampling with replacement technique that can be used to estimate the properties of a population from a sample.
Confusion matrix	A table or matrix that describes the ratio of correctly classified values with those that have been *confused* and classed under the wrong associated label.
Cross Validation	A technique used to assess statistical analyses, involving the random partitioning of data into different subsets, using one subset to train and the others to validate. This is repeated for a number of iterations defined by the analyst.
Data mining	The examination and study of databases to create new information.
Data science	A field of research that involves the application of scientific techniques to extract knowledge from data.

Elbow point	The point where no additional iterations or groups are able to produce significant changes in representation in within-group homogeneity or heterogeneity.
Epoch	The passing of an entire dataset through a neural network only once. This is an iterative process, which can be repeated multiple times as specified by the analyst.
False negative	A test which wrongly classifies an element as absent.
False positive	A test which wrongly classifies an element as present.
Feature engineering	The adjustment of hyperparameters and model architectures to ensure the best results.
Frequentist or Fisherian statistics	An approach to statistically calculating probabilities based on the frequency or proportion of an event based on sample data.
Gradient Descent	An optimisation algorithm that employs multiple mathematical concepts to define the steepest point of a graph where the least amount of error can be found based on a particular hyperparameter being optimised.
Heuristics	The ability to learn something for one's self.
Hyperparameters	A component of an algorithm whose value is defined before fitting the model to data. These are used to refine the learning process and optimise the performance of the algorithm.
Kappa	A coefficient (κ) that measures the degree of agreement of testing predictions in order to define the reliability of a fitted model.
K-means	A number that defines the number of centroids being defined in a test.
Loss	The error produced by a model when making a prediction. This can be measured by a variety of different functions.
Neural Networks	A combination of mathematical computations that mimic the structure and functionality of the human brain when processing data.
One-hot-encoding	The process where categorical variables are converted into binary form.
Optimisation	A means of finding the optimal performance of a model, thus minimising error defined by a loss function, through a combination of different hyperparameter settings and weights used for predictive computations.
Overfitting	An error produced when the model is fit too closely to training data, without the ability to generalise and thus more likely to make mistakes when exposed to new data.
Reinforcement Learning	Models that are trained on data based on the environment that they interact with, receiving a reward signal when correctly reacting to a problem solving situation.
Sensitivity	The ability to correctly identify the true positive.
Specificity	The ability to correctly identify the true negative.
Stochastic	The state of having a random distribution or pattern.
Supervised Learning	Models that are trained on data with labelled output values.
True negative	A test which correctly classifies an element as absent.
True positive	A test which correctly classifies an element as present.
Underfitting	An error produced when the model is not fit closely enough to training data, without the ability to capture underlying trends within the data.

Unsupervised Learning	Models that are trained on data with unlabelled output values, forcing the models to detect patterns within the data by themselves.
Weights	Numerical values that condition the signal being sent to the output of the model.

Appendix 2
Additional Graphs and Tables

Graphs and tables with supporting information for the Results chapter of this Master's Thesis.

Contents:

Cut Mark Number	Lighting Position	Length of Incision (mm)	Width of Incision (mm)	Magnification	Focus Range (Lens Height) (mm)		Number of Pixels		Number of Photos Taken per Tile	Number of Tiles	Tiling Time (Min)	Time Taken to Place Landmarks (Min)	Total Processing Time (Min)
					Sup.	Inf.	X	Y					
1	Ring/Coaxial	8.494	0.935	200	815.95	-1583.05	2411	9844	25	42	20:47	06:22	27:09
	Coaxial	8.494	0.935	200	815.95	-1583.05	2411	9844	25	42	20:50	04:59	25:49
	Left	8.494	0.935	200	815.94	-1583.05	2411	9844	25	39	19:22	04:59	24:21
	Right	8.494	0.935	200	815.95	-1583.05	2411	9844	25	39	19:10	03:17	22:27
2	Ring/Coaxial	11.552	1.455	100	3763.6	-470.7	2496	7175	30	27	13:31	07:52	21:23
	Coaxial	11.552	1.455	100	3763.6	-470.7	2496	7175	30	30	14:55	06:15	21:10
	Left	11.552	1.455	100	3763.6	-470.7	2496	7175	30	30	14:40	06:05	20:45
	Right	11.552	1.455	100	3763.6	-470.7	2496	7175	34	23	12:33	07:17	19:50
3	Ring/Coaxial	6.521	0.682	200	1020.9	-253.3	1731	7479	25	22	08:07	05:20	13:27
	Coaxial	6.521	0.682	200	1020.9	-253.3	1731	7479	25	23	05:32	07:18	12:50
	Left	6.521	0.682	200	1020.9	-253.3	1731	7479	25	23	08:39	05:09	13:12
	Right	6.521	0.682	200	1020.9	-253.3	1731	7479	25	30	08:04	05:30	13:34
4	Ring/Coaxial	9.000	1.197	150	1531.2	-841.8	2807	7960	25	34	13:24	07:12	20:36
	Coaxial	9.000	1.197	150	1531.2	-841.8	2807	7960	25	34	13:50	06:54	20:44
	Left	9.000	1.197	150	1531.2	-841.8	2807	7960	25	34	14:33	07:48	21:21
	Right	9.000	1.197	150	1531.2	-841.8	2807	7960	25	31	13:23	05:24	18:47

Table S1 – Table presenting the different variables taken when carrying out a 3 Dimensional Reconstruction using the HIROX KH-8700 Digital Microscope and the time taken to process each cut mark under the different lighting conditions.

Technique	System	Measuring Procedure	Classification	Portability	Full 3D Reconstruction Time (Min)	Operable Distance	Resolution	Cost (Eur)
Microscope	KH-8700	3D Digital Microscope	Active Sensor	Low	Aprox. 13.5	1-10 mm	0.15 - 0.01 µm	<100,000
Laser Scanner	David SLS-2	Structured Light	Active Sensor	Medium	>1	0.15 - 5.00 m	0.02 mm	3,000
Photogrammetry	Reflex + Macro Objective	Micro-photogrammetry	Passive Sensor	High	Aprox. 25	10 - 50 cm	0.02 mm	1,000

Table S2 – Comparisons between different reconstruction techniques. Updated from Maté-González et al. (2017c)

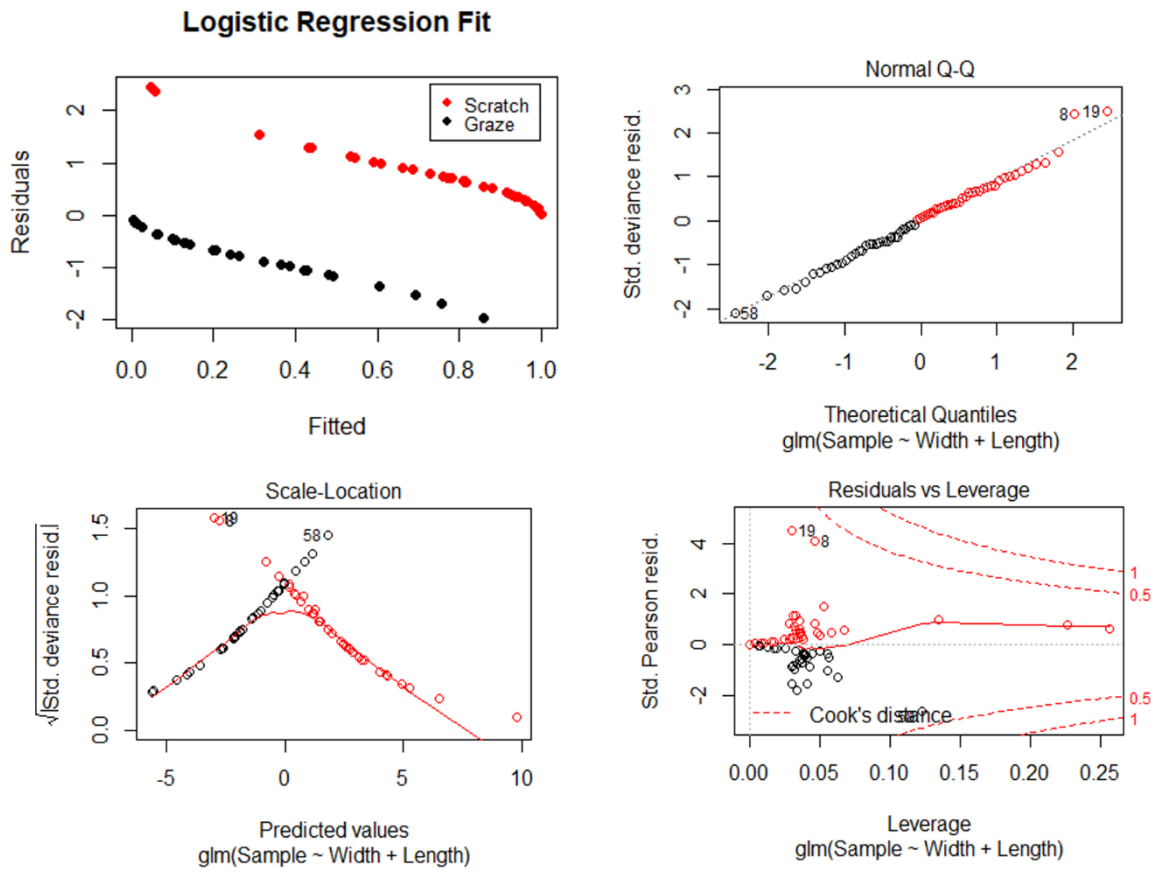


Figure S1 – Regression residual plots of final logistic regression ("glm") model used to distinguish scratches and grazes based on width and length variables. Top left – Residual plot against fitted values. Top right – Normal of residual plot describing the assumption that residuals are normally distributed. Bottom left – Scaled location measuring the root of standardised residuals. Bottom right – Standardised residuals plotted against leverage, assessing the effect of each data point on the fitted regression model.

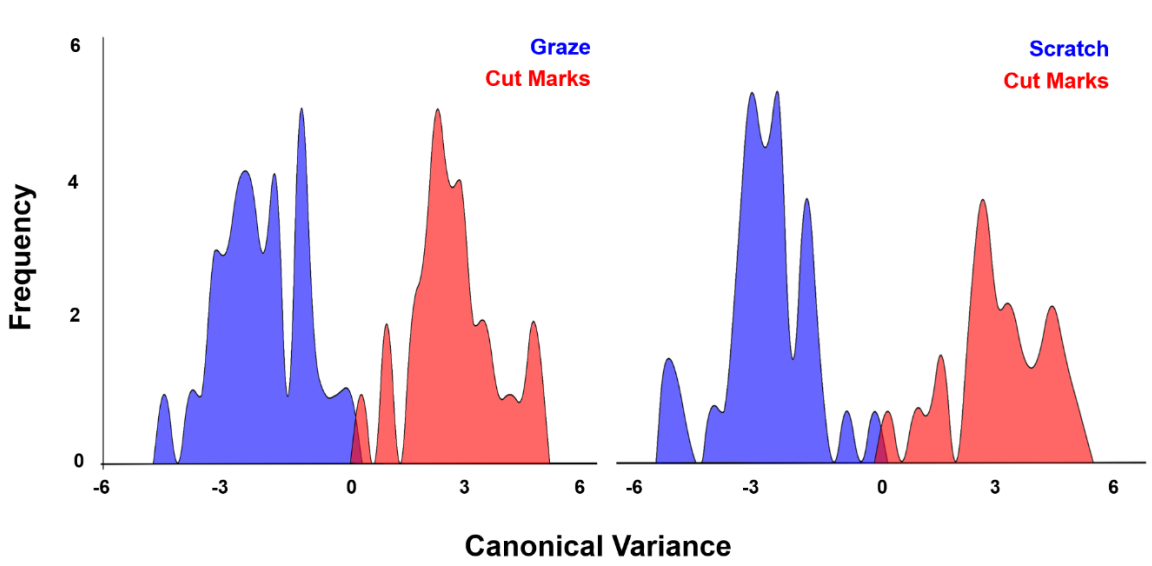


Figure S2 – CVA between cut marks and each type of trampling marks.

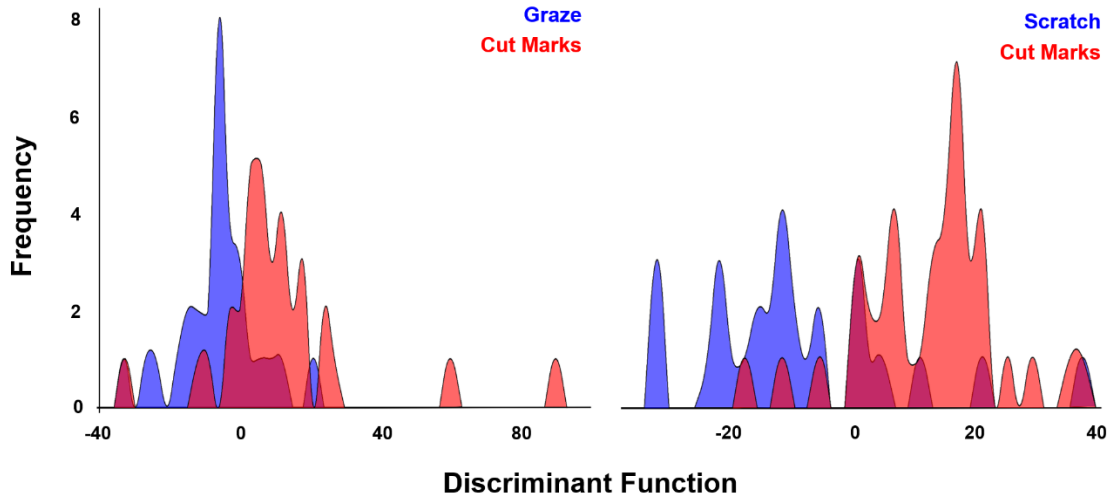


Figure S3 – Cross Validated Linear Discriminant Function graph to show the degree of separation between cut marks and different types of trampling marks using LDA techniques.

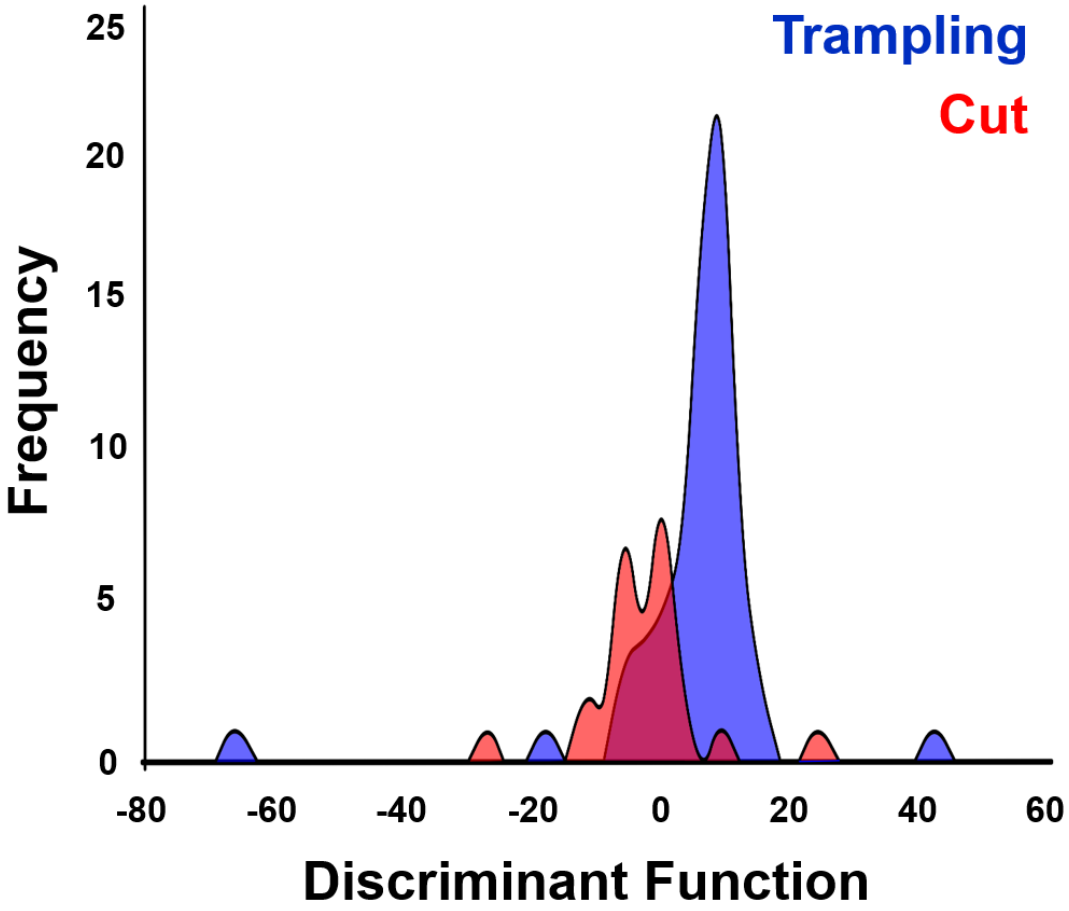


Figure S4 – Cross Validated Linear Discriminant Function graph to show the degree of separation between cut marks and trampling marks using LDA techniques.

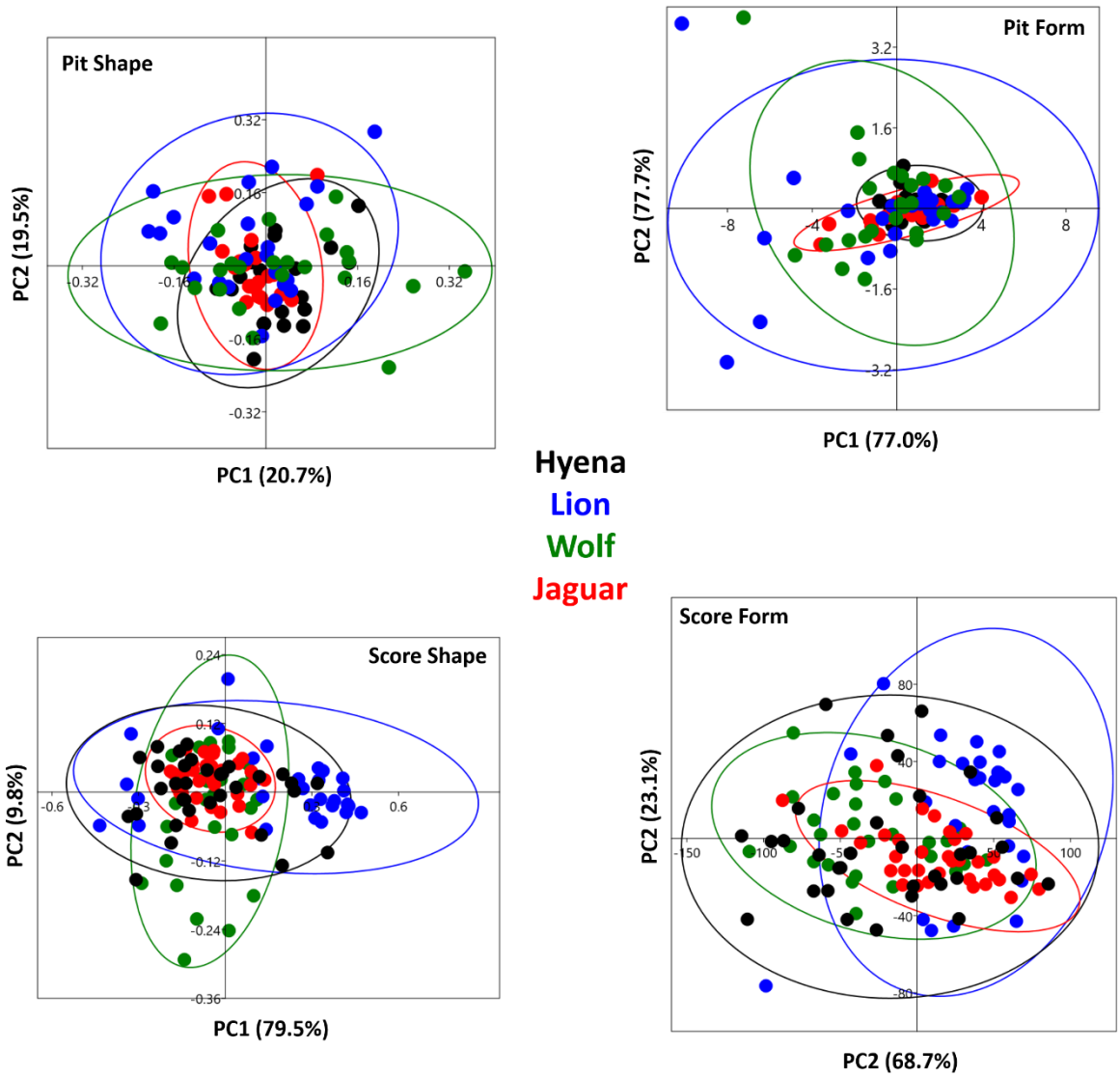


Figure S5 – PCA scatterplots describing variation in shape and form for different carnivores and different types of alterations

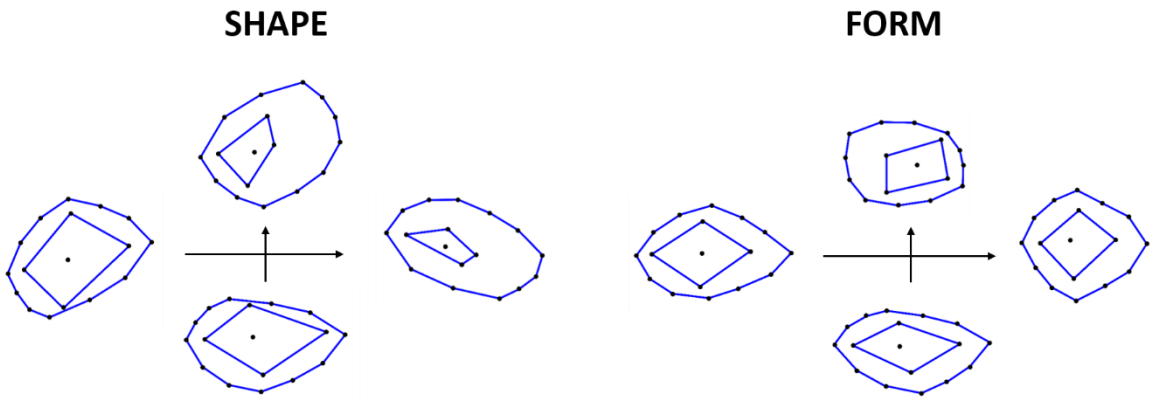


Figure S6 – Changes in shape and form of tooth pits across PC scores 1 (horizontal) and 2 (vertical) for both shape and form.

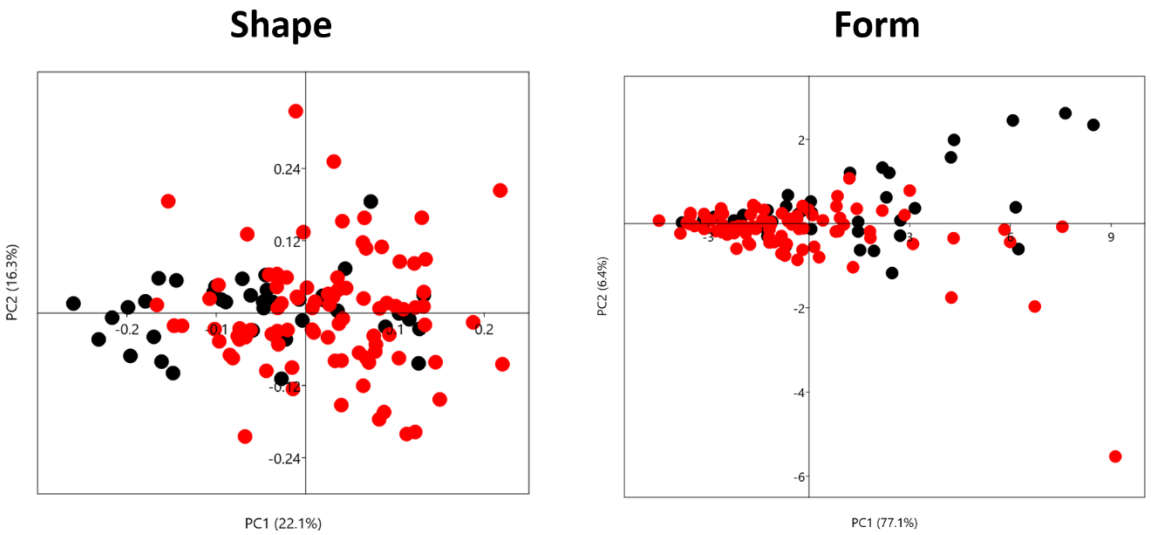


Figure S7 – PCA plots in shape and form comparing percussion marks (black) from tooth marks (red)

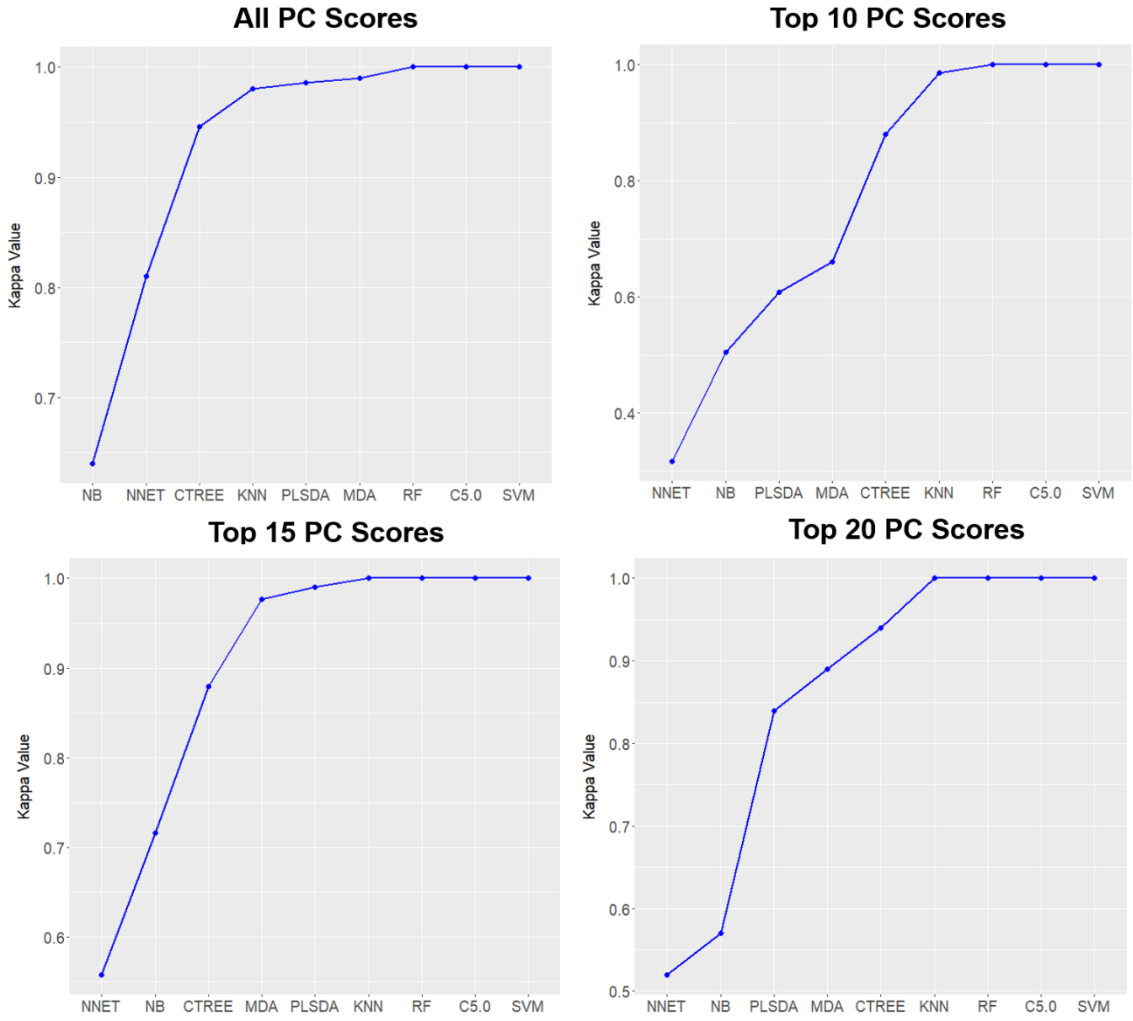


Figure S8 – Kappa curves assessing ML algorithm performance after tuning out poor performing PC scores for carnivore tooth pits in form space.

	Model Performance		95% Confidence Interval			
Model	**Kappa**	**Accuracy**	**Lower**	**Upper**	**Spec.**	**Sens.**
NNET	0.317	0.48	0.42	0.54	0.84	0.63
NB	0.50	0.63	0.57	0.68	0.88	0.63
PLSDA	0.61	0.71	0.65	0.757	0.90	0.70
MDA	0.66	0.75	0.69	0.79	0.91	0.75
CTREE	0.88	0.91	0.88	0.94	0.95	0.94
KNN	0.986	0.99	0.97	1	1	1
RF	1	1	0.99	1	1	1
C5.0	1	1	0.99	1	1	1
SVM	1	1	0.99	1	1	1

Table S3 – Comparisons of Machine Learning models trained on Top 10 scoring PC scores out of scores for carnivore tooth pits in form

	Model Performance		95% Confidence Interval			
Model	**Kappa**	**Accuracy**	**Lower**	**Upper**	**Spec.**	**Sens.**
NNET	0.56	0.67	0.61	0.72	0.90	0.79
NB	0.72	0.79	0.74	0.83	0.93	0.79
CTREE	0.88	0.91	0.87	0.94	0.97	0.91
MDA	0.98	0.98	0.96	0.99	0.99	0.98
PLSDA	0.99	0.99	0.98	0.99	1	1
KNN	1	1	0.99	1	1	1
RF	1	1	0.99	1	1	1
C5.0	1	1	0.99	1	1	1
SVM	1	1	0.99	1	1	1

Table S4 – Comparisons of Machine Learning models trained on Top 15 scoring PC scores out of scores for carnivore tooth pits in form

	Model Performance		95% Confidence Interval			
Model	**Kappa**	**Accuracy**	**Lower**	**Upper**	**Spec.**	**Sens.**
NNET	0.52	0.65	0.59	0.70	0.88	0.70
NB	0.57	0.68	0.62	0.73	0.89	0.67
PLSDA	0.84	0.88	0.91	1.00	0.96	0.88
MDA	0.89	0.92	0.89	0.95	0.97	0.92
CTREE	0.94	0.96	0.92	0.98	0.98	0.95
KNN	1	1	0.99	1	1	1
RF	1	1	0.99	1	1	1
C5.0	1	1	0.99	1	1	1
SVM	1	1	0.99	1	1	1

Table S5 – Comparisons of Machine Learning models trained on Top 20 scoring PC scores out of scores for carnivore tooth pits in form

		Miliseconds			Seconds		
		Min	**Mean**	**Max**	**Min**	**Mean**	**Max**
Linear Traces	**NNET**	52.24	100.20	171.04	0.05	0.10	0.17
	SVM	22.51	26.63	51.44	0.02	0.03	0.05
	KNN	2.49	3.01	5.67	0.00	0.00	0.01
	RF	41172.68	44457.79	53070.94	41.17	44.46	53.07
	MDA	22.50	29.58	119.73	0.02	0.03	0.12
	NB	2.24	2.63	3.92	0.00	0.00	0.00
	C5.0	34.04	40.36	89.75	0.03	0.04	0.09
	PLSDA	3547.75	3800.07	5023.12	3.55	3.80	5.02
	CTREE	8.18	12.38	250.84	0.01	0.01	0.25
Circular Depressions	**NNET**	216.08	256.91	296.41	0.22	0.26	0.30
	SVM	75.26	79.15	100.94	0.08	0.08	0.10
	KNN	12.49	13.36	19.70	0.01	0.01	0.02
	RF	192001.14	193674.09	201501.73	192.00	193.67	201.50
	MDA	60.40	68.34	91.05	0.06	0.07	0.09
	NB	14.59	16.16	24.73	0.01	0.02	0.02
	C5.0	153.73	163.68	188.71	0.15	0.16	0.19
	PLSDA	14916.20	15190.42	15813.13	14.92	15.19	15.81
	CTREE	44.68	47.96	62.97	0.04	0.05	0.06
Cross-Sections	**NNET**	99.57	177.18	226.85	0.10	0.18	0.23
	SVM	25.87	28.49	47.26	0.03	0.03	0.05
	KNN	3.22	3.58	6.61	0.00	0.00	0.01
	RF	79796.41	82344.51	108519.10	79.80	82.34	108.52
	MDA	37.04	42.06	56.74	0.04	0.04	0.06
	NB	4.26	4.81	16.50	0.00	0.00	0.02
	C5.0	46.67	49.99	62.54	0.05	0.05	0.06
	PLSDA	5926.33	6151.71	8813.44	5.93	6.15	8.81
	CTREE	14.49	16.55	26.75	0.01	0.02	0.03

Table S6 – Table presenting the time in miliseconds and seconds taken to train each ML model. Mean time recordings are calculated after 200 iterations using the *microbenchmark* function in R.

	Miliseconds		
	Min	**Mean**	**Max**
NNET	NNET	1.25	1.52
SVM	SVM	4.36	5.32
KNN	KNN	3.22	3.75
RF	RF	10.09	11.38
MDA	MDA	11.51	13.91
NB	NB	68.97	81.03
C5.0	C5.0	29.64	34.78
PLSDA	PLSDA	5.10	6.04
CTREE	CTREE	3.00	3.94

Table S7 – Table presenting the time in miliseconds taken to make a prediction using each ML model. Mean time recordings are calculated after 200 iterations using the *microbenchmark* function in R.

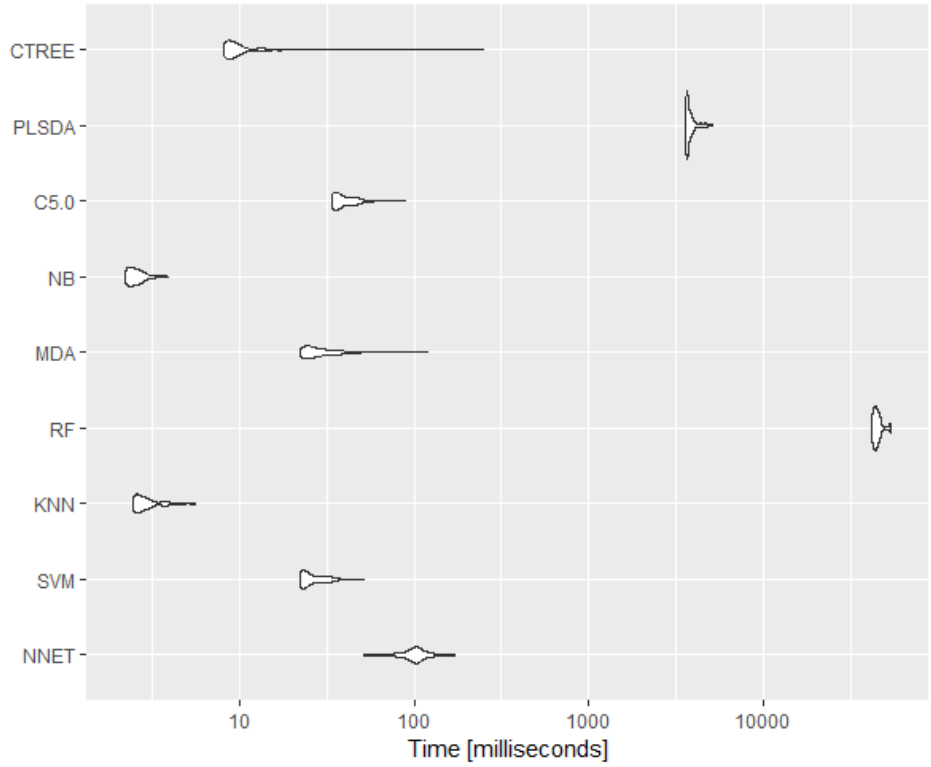


Figure S9 – An example of the time taken to train a ML model on GMM data processing Linear Incisions.

Algorithm	Input/Output		CPU Utilisation (%)						CPU Speed (GHz)					
	In	Out	Min	Mean	SD	Lower CI	Upper CI	Max	Min	Mean	SD	Lower CI	Upper CI	Max
NNET	10	2	16	28.67	7.45	26.43	30.91	43	2.74	2.89	0.06	2.87	2.91	2.99
	51	2	11	26.97	9.58	24.09	29.85	46	2.71	2.89	0.04	2.88	2.90	2.92
	51	4	18	31.63	9.65	28.74	34.53	64	2.81	2.89	0.03	2.88	2.90	2.94
	14	4	11	21.97	7.26	19.79	24.15	35	2.80	2.88	0.03	2.87	2.89	2.94
SVM	10	2	14	23.93	5.70	22.22	25.64	39	2.75	2.91	0.06	2.90	2.92	2.98
	51	2	10	25.70	8.89	23.03	28.37	49	2.84	2.88	0.02	2.87	2.89	2.92
	51	4	12	24.67	7.41	22.44	26.90	41	2.76	2.88	0.03	2.87	2.89	2.92
	14	4	14	20.67	4.74	19.25	22.09	31	2.79	2.90	0.04	2.89	2.91	2.94
KNN	10	2	12	19.57	6.40	17.65	21.49	42	2.63	2.83	0.06	2.81	2.85	2.91
	51	2	9	16.20	5.27	14.62	17.78	33	2.81	2.87	0.02	2.86	2.88	2.90
	51	4	10	18.00	4.80	16.65	19.35	30	2.81	2.86	0.03	2.85	2.87	2.92
	14	4	10	17.20	7.20	15.04	19.36	39	2.80	2.86	0.02	2.85	2.87	2.90
RF	10	2	38	39.07	1.46	38.63	39.51	43	2.96	3.07	0.04	3.06	3.08	3.13
	51	2	33	39.50	4.08	38.27	40.73	57	2.94	3.08	0.04	3.07	3.09	3.13
	51	4	37	39.53	3.97	38.34	40.72	57	2.90	3.08	0.05	3.07	3.10	3.13
	14	4	25	37.90	3.21	36.94	38.86	47	3.03	3.07	0.02	3.06	3.08	3.12
MDA	10	2	16	24.60	6.84	22.55	26.65	41	2.74	2.88	0.04	2.87	2.89	2.95
	51	2	12	22.93	6.27	21.05	24.81	40	2.62	2.85	0.06	2.83	2.87	2.92
	51	4	11	22.63	7.18	20.47	24.79	39	2.73	2.87	0.04	2.86	2.88	2.91
	14	4	11	17.23	5.37	15.62	18.84	30	2.80	2.87	0.04	2.86	2.88	3.04
NB	10	2	11	22.70	6.05	20.88	24.52	39	2.79	2.89	0.04	2.88	2.90	2.94
	51	2	21	34.53	7.26	32.35	36.71	56	2.76	2.85	0.03	2.84	2.86	2.90
	51	4	14	27.60	7.38	25.38	29.82	40	2.80	2.89	0.02	2.88	2.90	2.95
	14	4	18	28.93	8.20	26.47	31.39	46	2.85	2.89	0.02	2.88	2.90	2.94
C5.0	10	2	15	22.17	5.14	20.63	23.71	36	2.55	2.91	0.02	2.90	2.92	2.95
	51	2	15	26.23	8.22	23.76	28.70	47	2.74	2.86	0.03	2.85	2.87	2.90
	51	4	13	20.30	6.94	18.22	22.38	36	2.67	2.87	0.04	2.86	2.88	2.94
	14	4	15	22.17	5.14	20.63	23.71	33	2.85	2.91	0.02	2.90	2.92	2.94
PLSDA	10	2	27	46.57	5.96	44.78	48.36	55	2.92	3.06	0.05	3.04	3.08	3.15
	51	2	35	49.20	5.67	47.50	50.90	64	3.01	3.08	0.02	3.07	3.09	3.12
	51	4	18	35.90	7.17	33.75	38.05	48	3.02	3.07	0.03	3.06	3.08	3.14
	14	4	21	46.70	7.11	44.57	48.83	61	2.72	2.87	0.04	2.86	2.88	3.01
CTREE	10	2	14	24.43	6.67	22.40	26.46	37	2.79	2.85	0.03	2.84	2.86	2.89
	51	2	11	20.20	7.47	17.77	22.63	43	2.76	2.85	0.03	2.84	2.86	2.92
	51	4	10	18.20	5.74	16.48	19.92	34	2.78	2.89	0.03	2.88	2.90	2.93
	14	4	16	24.57	8.11	22.14	27.00	43	2.74	2.87	0.04	2.86	2.88	2.99

Table S8 – Comparison of CPU processor performance during training of different ML models. Input refers to the number of dependencies inserted into each model during training and the Output refers to the number of classifier labels the model is trained to reach.

Density Layer Configuration	Kernel Init.	Activation Parameters	Dropout parameters	Weight Regulizer	Loss	Optimizer	Metrics	Valid. split	Epochs	Batch Size	Accuracy
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	150	10	87.4
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	100	10	88
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	100	10	90.3
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	100	15	88
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	200	15	88.9
10, 5, 1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	200	15	89
10,5,2,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	200	15	70.3
10,5,2,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	30	300	20	88
10,5,2,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	mse	Accuracy	30	300	20	88.4
10,5,2,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	mse	Accuracy	30	200	20	87.9
10,5,2,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	mse	Accuracy	30	300	20	87.9
10,8,3,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	200	15	88
10,8,3,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	300	20	89
10,8,3,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	300	20	89
10,8,3,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	300	50	88.4
10,8,3,1	Uniform	Relu, Sigmoid	NONE	NONE	Binary Crossentropy	Adam	Accuracy	20	500	50	88
10,8,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	20	400	50	89.1
10,5,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	20	400	50	88
10,5,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	20	450	50	70.3
10,5,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	20	450	100	91.4
10,8,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	20	1000	100	92.4
10,8,3,1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	RMSprop	Accuracy	20	1000	100	89.1
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	NONE	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	100	96.1
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	100	93.9
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	2000	100	96.4
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	100	150	92.7
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	500	200	89.7
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	NONE	Binary Crossentropy	Adam	Accuracy	30	500	200	87.1
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5	NONE	Binary Crossentropy	Adam	Accuracy	30	1000	200	87.1

10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	2000	100	96.1
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	2000	150	96.1
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	2000	150	95.43
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1250	150	95.43
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	95.86
10, 20, 10, 5, 1	Uniform	Relu, Sigmoid	0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	95.86
10, 20, 10, 8, 5, 1	Uniform	Relu, Sigmoid	0.5, 0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	95.14
10, 20, 10, 8, 5, 1	Uniform	Relu, Sigmoid	0.5, 0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	95.14
10, 20, 10, 8, 5, 1	Uniform	Relu, Sigmoid	0.5, 0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1500	200	91.14
10, 20, 8, 4	Uniform	Relu, Sigmoid	0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	94.43
10, 20, 8, 4	Uniform	Relu, Sigmoid	0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	800	150	94.57
10, 20, 8, 4	Uniform	Relu, Sigmoid	0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	500	150	91.71
10, 20, 8, 4	Uniform	Relu, Sigmoid	0.5, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	500	300	91.43
10, 20, 12, 6, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	91.71
10, 20, 12, 6, 1	Uniform	Relu, Sigmoid	0.5, 0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	90.14
10, 20, 10, 6, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	97.14
10, 20, 10, 6, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	1000	150	95.86
10, 20, 10, 6, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	900	32	92.29
10, 20, 10, 6, 1	Uniform	Relu, Sigmoid	0.8	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	900	64	95.86
10, 20, 10, 6, 1	Uniform	Relu, Sigmoid	0.8, 0.5	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	900	64	95.86
10, 20, 10, 6, 1	None	Relu, Sigmoid	0, 0, 0, 0.5, 0, 0	Kernel Constraint	Binary Crossentropy	Adam	Accuracy	30	900	64	100

Table S9 – Table presenting the different hyperparameters tested when training a DNN on cut mark and trampling mark data.

Model	Time (s)
Cut vs Trampling	10.87
Carnivore Pits	5.41
Percussion vs Tooth	8.01
Taxa Scores	37.94

Table S10 – Time taken to train each DNN Model

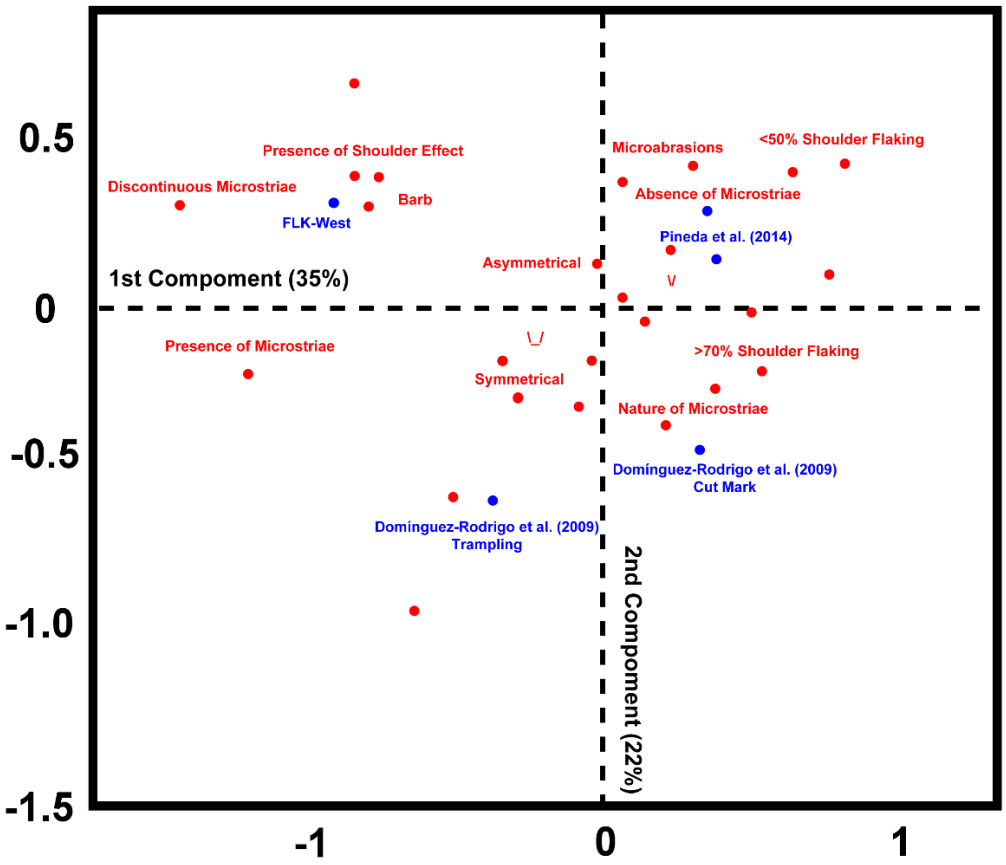


Figure S10 – Multiple Correspondence Analysis of qualitative features observed in linear incisions at FLK-West.

SVM		RF		C5.0		DNN	
Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)
Cut Mark	94.29	Trampling	67.20	Cut Mark	50.29	Trampling	96.74
Trampling	100.00	Trampling	74.80	Cut Mark	50.29	Trampling	99.78
Trampling	75.18	Cut Mark	52.40	Cut Mark	50.29	Trampling	99.25
Trampling	97.94	Trampling	81.60	Cut Mark	50.29	Trampling	99.99
Cut Mark	86.15	Trampling	68.40	Cut Mark	50.29	Cut Mark	98.76
Cut Mark	99.95	Cut Mark	77.40	Cut Mark	50.29	Cut Mark	95.69
Cut Mark	99.98	Cut Mark	66.00	Cut Mark	50.29	Cut Mark	98.65
Trampling	65.02	Cut Mark	68.20	Cut Mark	50.29	Trampling	82.00
Trampling	99.93	Trampling	74.00	Cut Mark	50.29	Trampling	99.70
Trampling	88.50	Trampling	62.40	Cut Mark	50.29	Trampling	99.95
Cut Mark	99.93	Trampling	57.60	Cut Mark	50.29	Cut Mark	62.17
Cut Mark	99.94	Cut Mark	59.00	Cut Mark	50.29	Cut Mark	98.50
Cut Mark	99.97	Cut Mark	62.60	Cut Mark	50.29	Cut Mark	94.31
Cut Mark	97.31	Cut Mark	72.80	Cut Mark	50.29	Cut Mark	96.85
Trampling	99.87	Cut Mark	54.40	Cut Mark	50.29	Trampling	99.93
Cut Mark	99.93	Cut Mark	59.20	Cut Mark	50.29	Cut Mark	94.43
Cut Mark	99.99	Trampling	57.40	Cut Mark	50.29	Cut Mark	98.10
Cut Mark	99.47	Cut Mark	83.80	Cut Mark	50.29	Trampling	96.89
Cut Mark	99.93	Cut Mark	54.20	Cut Mark	50.29	Trampling	73.88
Trampling	76.53	Cut Mark	60.60	Cut Mark	50.29	Cut Mark	100.00
Cut Mark	99.85	Cut Mark	88.00	Cut Mark	50.29	Cut Mark	98.57
Trampling	99.99	Trampling	57.00	Cut Mark	50.29	Trampling	99.88
Trampling	98.97	Cut Mark	53.20	Cut Mark	50.29	Cut Mark	99.27
Cut Mark	99.19	Trampling	60.00	Cut Mark	50.29	Trampling	98.91

Table S11 – Initial classification of linear marks from FLK-West using all 3 ML models and the 1 DL model.

SVM		RF		C5.0		DNN	
Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)
Graze	99.99	Graze	69.20	Scratch	51.28	Graze	99.94
Graze	100.00	Graze	86.80	Scratch	51.28	Graze	100.00
Graze	100.00	Graze	80.60	Scratch	51.28	Graze	100.00
Graze	100.00	Graze	75.20	Scratch	51.28	Graze	100.00
Scratch	91.47	Scratch	65.40	Scratch	51.28	Scratch	80.96
Graze	100.00	Graze	83.80	Scratch	51.28	Graze	99.99
Graze	100.00	Graze	80.20	Scratch	51.28	Graze	99.99
Graze	100.00	Graze	64.80	Scratch	51.28	Graze	99.99
Graze	100.00	Graze	84.40	Scratch	51.28	Graze	100.00

Table S12 – Initial classification of trampling marks from FLK-West using all 3 ML models and the 1 DL model.

SVM		RF		C5.0		DNN	
Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)
Tooth	83.29	Tooth	50.40	Tooth	65.71	Percussion	98.79
Tooth	61.57	Tooth	68.40	Tooth	65.71	Tooth	100.00
Tooth	83.42	Tooth	69.20	Tooth	65.71	Tooth	100.00
Tooth	82.07	Tooth	68.40	Tooth	65.71	Tooth	100.00

Table S13 – Initial classification of circular depressions from FLK-West using all 3 ML models and the 1 DL model.

SVM		RF		C5.0		DNN	
Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)	Class	Confidence (%)
Hyena	78.09	Lion	59.4	Lion	57.29	Lion	99.99
Hyena	65.01	Lion	65.2	Lion	57.29	Lion	85.11
Hyena	79.97	Lion	60	Lion	57.29	Lion	99.67

Table S14 – Initial classification of carnivore taxa using tooth pits from FLK-West using all 3 ML models and the 1 DL model.

Appendix 3
Restauration

Supporting figures for the Discussion chapter of this Master's Thesis. Two reports are included containing chemical data obtained using the ESEM.

Contents:

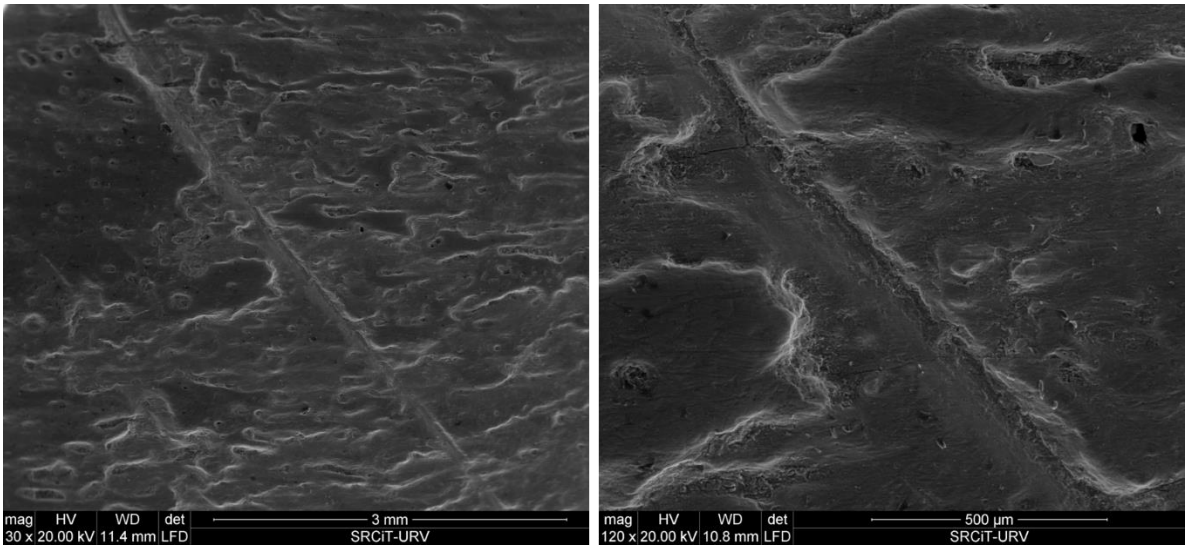


Figure S11 – SEM photos of consolidated linear incisions from Valparadís. Note how in the photo to the right, small fissures and cracks can be seen on the bone's surface, indicating its fragile nature and thus requiring consolidation treatment, however, also note how the consolidant covers most of the bone, making the perception of its features and depth much more difficult

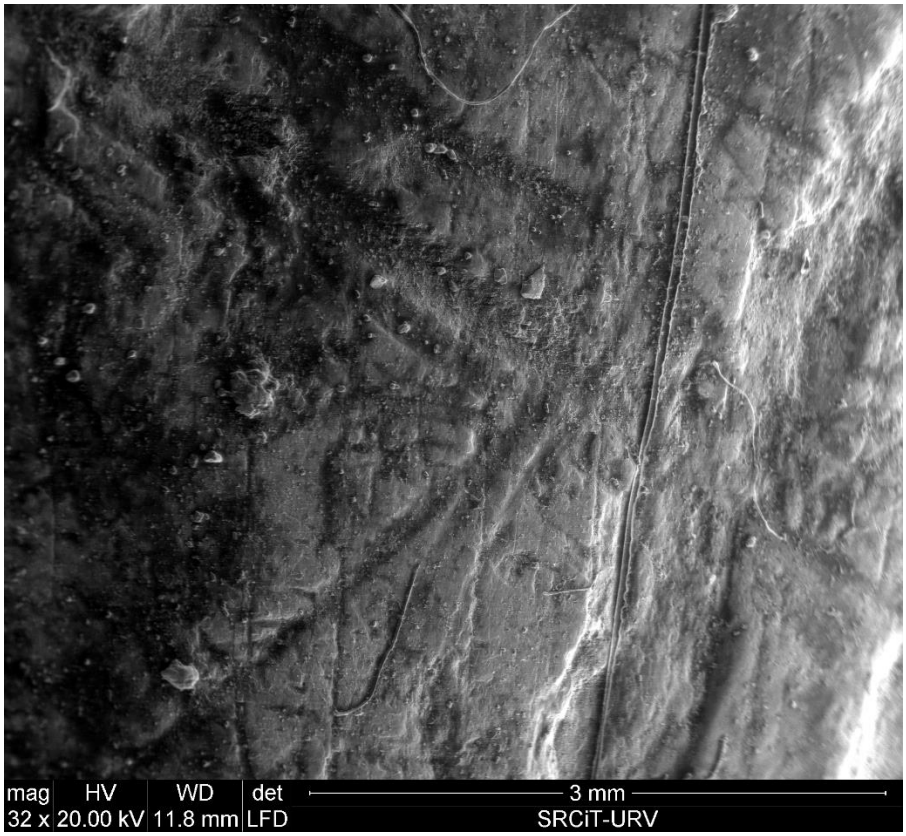
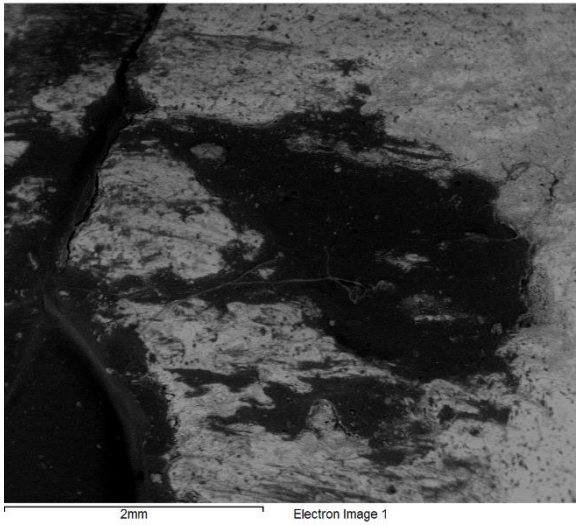
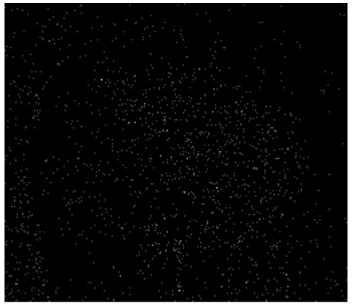
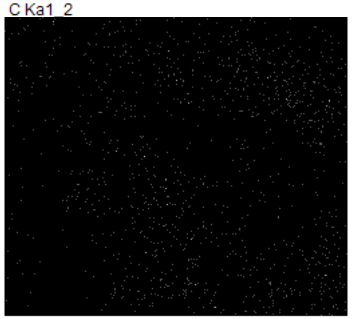


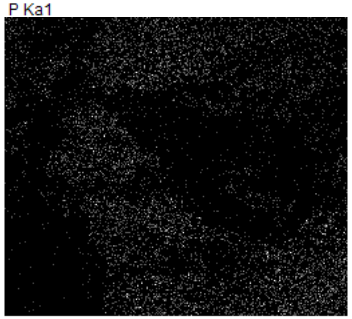
Figure S12 – SEM photos of linear Incisions across the surface of a bone from Valparadís. This bone has been consolidated, yet to a lesser degree than others presented within this study. Nevertheless, the consolidant has caused a series of fibers and dirt to become stuck to the bone's surface, making it hard to analyse the traces present.

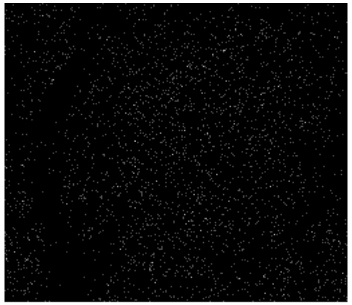


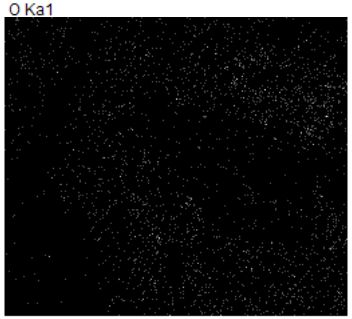

C Ka1_2


Al Ka1


P Ka1


Ca Ka1

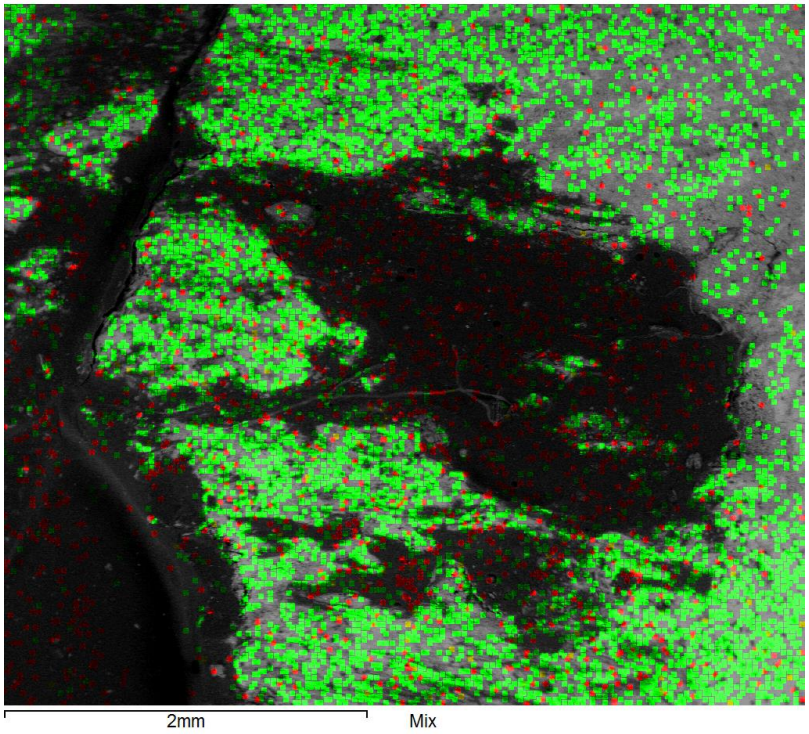

O Ka1


Si Ka1


K Ka1


Fe Ka1

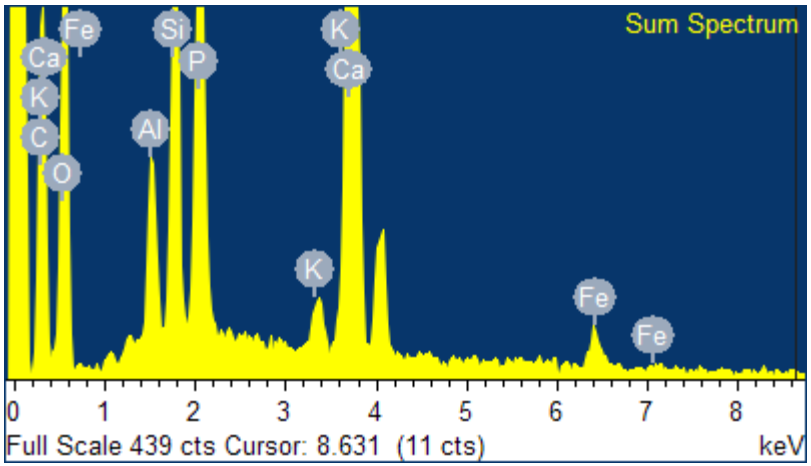
Report 1 - Chemical study of a consolidated bone's surface from Valparadís. Note how the presence of Paraloid B-72 covers the circular depression, making it hard to study and classify. Different analysis of chemical compositions indicate that organic material cannot be detected where the consolidant is present, thus showing how the electrons are unable to reach the cortical surface or the bottom of the mark. The additional presence of elements such as Silicon indicate the presence of dirt within the mark as well.



Processing option : All elements analyzed (Normalised)
Number of iterations = 4

Standard :
C CaCO3 1-Jun-1999 12:00 AM
O SiO2 1-Jun-1999 12:00 AM
Al Al2O3 1-Jun-1999 12:00 AM
Si SiO2 1-Jun-1999 12:00 AM
P GaP 1-Jun-1999 12:00 AM
K MAD-10 Feldspar 1-Jun-1999 12:00 AM
Ca Wollastonite 1-Jun-1999 12:00 AM
Fe Fe 1-Jun-1999 12:00 AM

Element	Weight%	Atomic%
C K	28.77	40.77
O K	42.93	45.67
Al K	1.60	1.01
Si K	3.43	2.08
P K	5.87	3.22
K K	0.68	0.29
Ca K	15.44	6.56
Fe K	1.30	0.40
Totals	100.00	



Comment: C-red, Ca-green

Figure S13 – Detailed photo of consolidated linear incision from Valparadís. Not how the consolidant creates a reflective surface as well as multiple bubbles that would consequently affect the quality of any 3D reconstruction

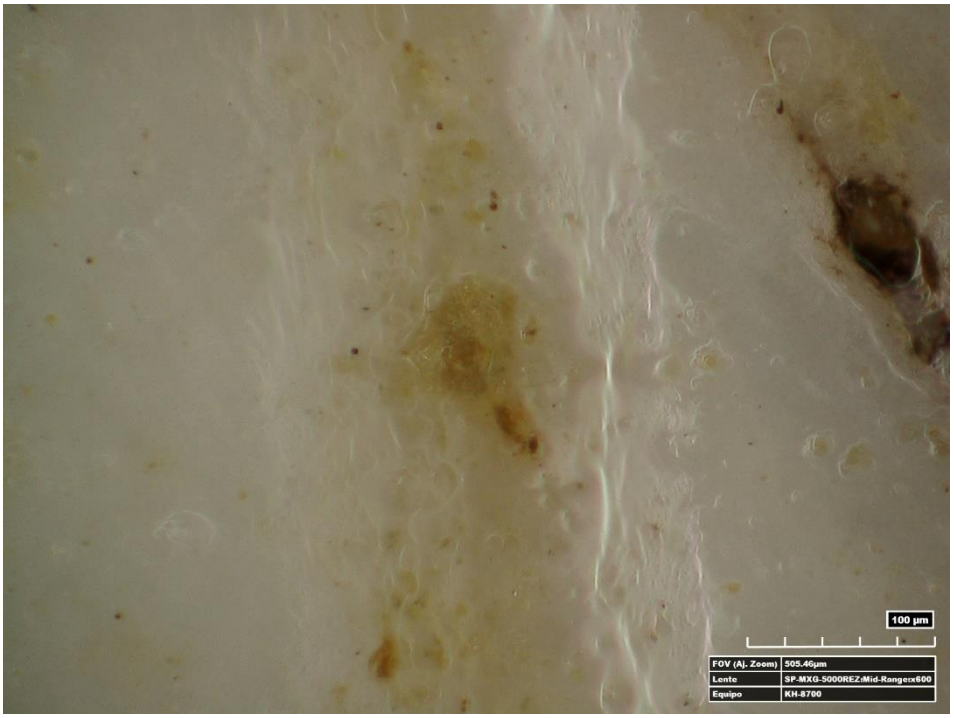


Figure S14 – Detailed photo of consolidated linear incision from Valparadís. Note how particles of dirt hang in suspension within the consolodiant

Figure S15 – Detailed photo of consolidated linear incision from Valparadís. Note the bubbles and reflective nature of the consolidant

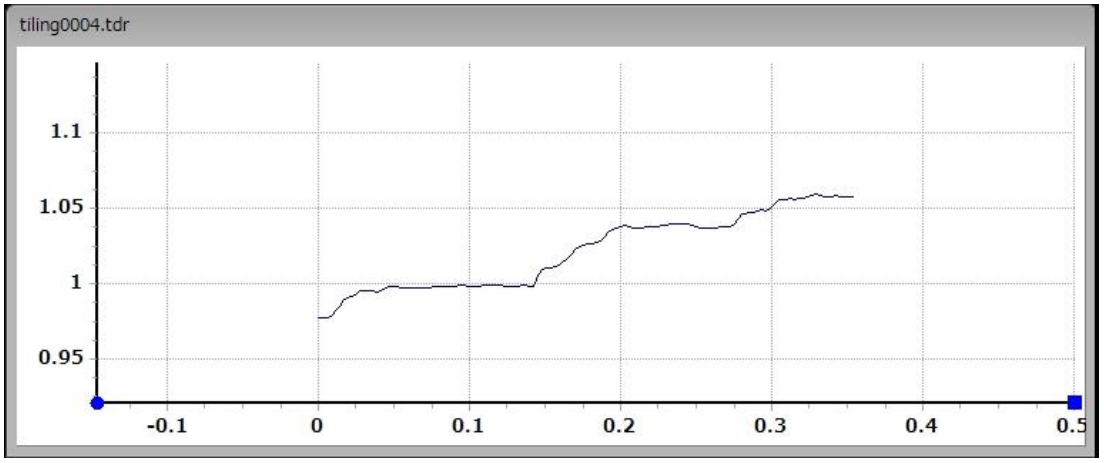


Figure S16 – Cross section obtained from the linear incision observed in Figure 5.

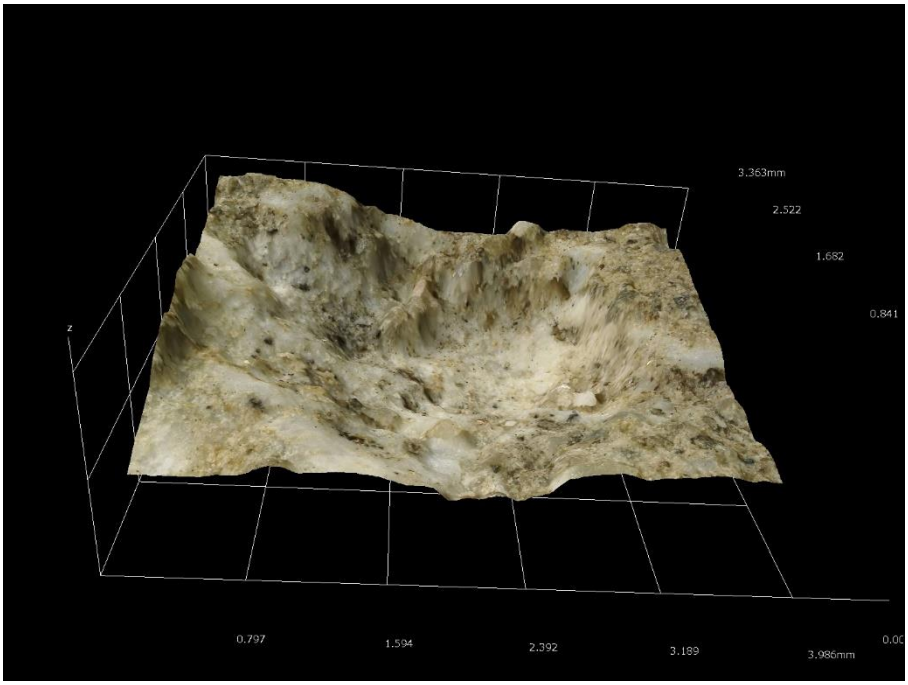


Figure S17 – 3D reconstruction of a circular depression from Valparadís. The presence of dirt prevents the analyst from being able to locate and document any landmark data for this particular mark, requiring technical assistance and cleaning before the mark can be analysed

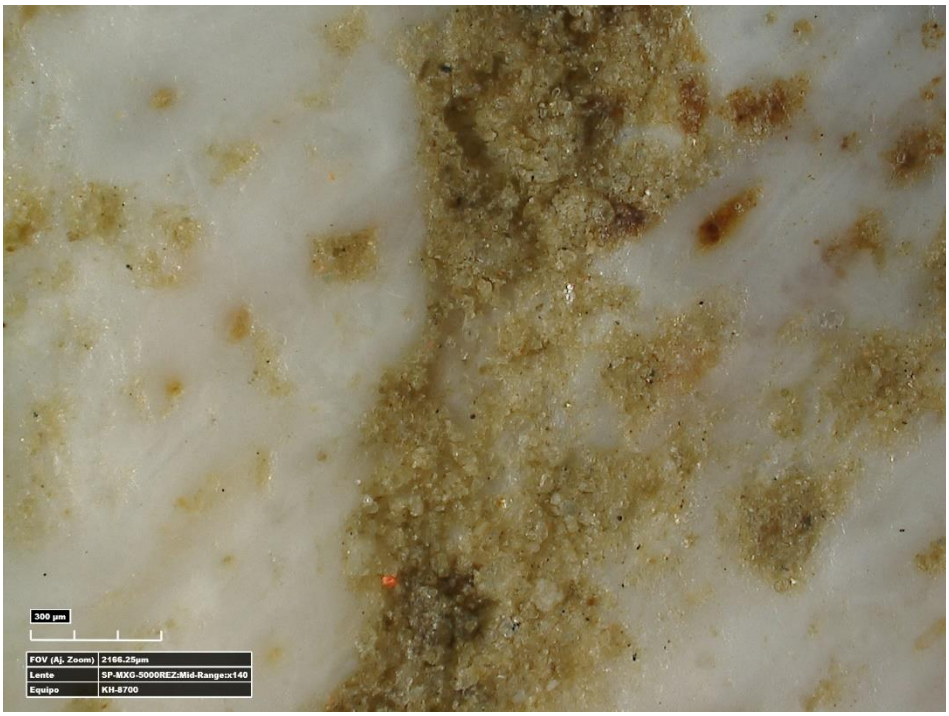


Figure S18 – Photo of a tooth score from Valparadís. The presence of dirt prevents the analyst from being able to locate and document any landmark data for this particular mark, requiring technical assistance and cleaning before the mark can be analysed

Appendix 4
Publication Number 1

The first published article produced with data from this Master's Thesis.

Courtenay, L.A.; Yravedra, J.; Huguet, R.; Ollé, A.; Aramendi, J.; Maté-González, M.Á.; González-Aguilera, D. (2018b) New Taphonomic Advances in 3D Digital Microscopy: a Morphological Characterisation of Trampling Marks, *Quaternary International*. DOI: 10.1016/j.quaint.2018.12.019

Quaternary International xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Quaternary International

journal homepage: www.elsevier.com/locate/quaint

New taphonomic advances in 3D digital microscopy: A morphological characterisation of trampling marks

Lloyd A. Courtenay[a,b,c,*], José Yravedra[c,d], Rosa Huguet[b,a,e], Andreu Ollé[b,a], Julia Aramendi[c,d], Miguel Ángel Maté-González[f], Diego González-Aguilera[f]

[a] Àrea de Prehistòria, Universitat Rovira i Virgili (URV), Avinguda de Catalunya 35, 43002, Tarragona, Spain
[b] Institut de Paleoecologia Humana i Evolució Social (IPHES). Zona Educacional, Campus Sescelades URV (Edifici W3) E3, 43700, Tarragona, Spain
[c] Department of Prehistory, Complutense University, Prof. Aranguren s/n, 28040, Madrid, Spain
[d] IDEA (Institute of Evolution in Africa), Covarrubias 36, 28010, Madrid, Spain
[e] Unit Associated to CSIC, Departamento de Paleobiologia, Museo de Ciencias Naturales, calle José Gutiérrez Abascal, s/n, 28006, Madrid, Spain
[f] Department of Cartography and Terrain Engineering, Polytechnic School of Avila, University of Salamanca, Hornos Caleros 50, 05003, Avila, Spain

ARTICLE INFO

Keywords:
Taphonomy
Microscopy
Bone surface modifications (BSM)
Geometric morphometrics

ABSTRACT

The concept of equifinality has become one of the greatest difficulties in the field of taphonomy However, new advances in technology have diminished the margins of error in the interpretation of archaeological sites. The use of multivariate statistics and the most recent advances in microscopic analysis of Bone Surface Modifications (BSMs) have enable a less subjective interpretation of site formation processes. Nevertheless, this broader range of methodological approaches also presents some problems. The capacity of laser scanners in processing inconspicuous and superficial cortical alterations, such as trampling marks, has proven to be problematic. This study presents a new advance towards resolving this problem through the use of the HIROX KH-8700 Digital Microscope, whereby detailed digital three-dimensional (3D) reconstructions are able to pick up such minute BSMs. Through the statistical comparison of the David Laser scanner and the HIROX KH-8700 Digital Microscope, this paper contributes to our understanding of said equipment, followed by a significant advance in the characterisation of superficial BSMs. The combination of advanced microscopy and the application of geometric morphometrics highlights a morphological differentiation between two different types of trampling marks, hereby named scratch and graze trampling marks.

1. Introduction

In recent years taphonomy has become a vital analytical tool for studying paleoanthropological sites. Taphonomy has provided a strong empirical background on which the interpretation of a site can be constructed, contributing to some of the most significant archaeological debates, such as the 'Hunter-Scavenger' debate (Binford, 1981; Blumenschine, 1986, 1995; Domínguez-Rodrigo, 1997; Domínguez-Rodrigo & Barba, 2006, 2007; Blumenschine et al., 2007; Domínguez-Rodrigo et al., 2007).

To many analysts, however, difficulties lie in two major concepts; analogy and uniformitarianism. Taphonomy relies heavily on the concept of uniformitarianism and actualism (Hutton, 1794; Playfair, 1802; Lyell et al., 1830; Whewell, 1847) to construct a hypothesis. Uniformitarianism is, in turn, highly conditioned by multiple factors relying on analogy for theoretical support (Bunge, 1981). In order to combine these concepts in the construction of theoretical frames of reference, experimentation plays a key role in order to provide a middle range theory that aids in our understanding of a site (Merton, 1967; Binford, 1967, 1968; 1981; Gifford-Gonzalez, 1991). Nevertheless, the incorrect use of analogy can condition our reconstruction of the past. Subjectivity plays a major role in this, strongly conditioned by the equifinality present in the taphonomic, paleoanthropological and archaeological record (Domínguez-Rodrigo et al., 2017). Equifinality is a conditioning factor that causes the misinterpretation of different taphonomic traces that can be product of the same agent, or *vice versa*. In recent years, the conflictive equifinality present between trampling and cut marks has become increasingly apparent, causing the misclassification of Bone Surface Modifications (BSMs) (Sahle et al., 2017), as well as the erroneous interpretation of sites (McPherron et al., 2010; Thompson et al., 2015).

Since their initial introduction in scientific literature, trampling

* Corresponding author. Àrea de Prehistòria, Universitat Rovira i Virgili (URV), Avinguda de Catalunya 35, 43002, Tarragona, Spain.
E-mail address: ladc1995@gmail.com (L.A. Courtenay).

https://doi.org/10.1016/j.quaint.2018.12.019
Received 25 October 2018; Received in revised form 17 December 2018; Accepted 19 December 2018
1040-6182/

marks have come to be defined as superficial, irregular and "flat bottomed" linear traces, produced by sedimentary abrasion (Brain, 1967; Behrensmeyer, 1978; Andrews and Cook, 1985). In the 1980's, issues with these definitions, however, led authors in disagreement as to the criteria used in defining naturally produced and anthropic linear striations (Fiorillo, 1984; Oliver, 1984; Andrews and Cook, 1985; Behrensmeyer et al., 1986). From this lack of supporting criteria, many analysts were led to highlight the possible equifinality present in the taphonomic register (Olsen and Shipman, 1988). While more in depth analysis of these traces were able to highlight the importance of different distinguishing variables (Domínguez-Rodrigo et al., 2009), analysts have still found issues using these qualitative variables, providing demand for more empirically objective data (Domínguez-Rodrigo et al., 2017).

Microscopy applied to taphonomy has played a key role in overcoming these issues. In the 1980's a great deal of studies revolved around the use of Scanning Electron Microscopes (SEM) (Potts and Shipman, 1981; Shipman, 1981; Shipman and Rose, 1983; Shipman et al., 1984a, b; Andrews and Cook, 1985; Behrensmeyer et al., 1986; Cook, 1986; Olsen, 1988; Olsen and Shipman, 1988), which took advantage of the high resolution and visual perception of texture. Combined with the introduction of statistical multivariate analyses in archaeology (Domínguez-Rodrigo et al., 2009; de Juana et al., 2010; Moclán et al., 2018), in recent years SEMs have proven to be a considerable tool in empirically processing the taphonomic register (Pineda et al., 2014). Come the 21st century, confocal microscopy (Archer and Braun, 2013; Pante et al., 2017; Otárolla-Castillo et al., 2017; Gümrükçu and Pante, 2018), high resolution optical microscopes (Bello and Soligo, 2008; Bello et al., 2009, 2016, 2013; Bello, 2011), as well as micro-photogrammetric three-dimensional (3D) reconstructions (Maté-González et al., 2015, 2016; Aramendi et al., 2017; Arriaza et al., 2017; Maté-González et al., 2017a, Yravedra et al., 2017a, b, c) have also been key tools in the development of analytical methods.

Further noteworthy developments are based on the analysis of two-dimensional (2D) derived information from 3D models, such as that of a cut mark's cross-section, via different biometric (Bello and Soligo, 2008; Bello et al., 2009, 2013; Bello, 2011) and geometric morphometric (Maté-González et al., 2015, 2016, 2017a) approaches. Multiple studies have also applied these concepts to BSMs produced by animals (Arriaza et al., 2017; Yravedra et al., 2017a). With the introduction of laser scanners, these approaches were then developed via 3D landmark models to further analyse cut marks (Courtenay et al., 2017) as well as other BSMs such as anthropic percussion pits (Yravedra et al., 2018) and carnivore tooth pits (Aramendi et al., 2017).

With such a broad range of techniques, methodological studies, comparisons and debates have been fundamental in contextualising and determining the value of these different methodological approaches. Work by Maté-González et al. (2017b, c), for example, have statistically analysed the reliability of different approaches, arguing the validity of the method. While some doubt has been cast over 2D landmark models (Otárolla-Castillo et al., 2017), further statistical data have strongly supported the reliability of these methods in taphonomic analysis (Courtenay et al., 2018).

The outcome of these methodological debates highlights the potential of laser scanners, such as the DAVID Structured-Light SLS-2 Scanner, as a considerably strong tool for future investigation in this field. The problem with such equipment lies in the lack of resolution when capturing superficial and inconspicuous taphonomic traces (Courtenay et al., 2017; Maté-González et al., 2017b, c), such as trampling marks – a problem that has become increasingly apparent in the taphonomic study of some archaeological and paleoanthropological sites (Domínguez-Rodrigo et al., 2011; Domínguez-Rodrigo and Alcalá, 2016; and further research pending publication). The objectives of this paper are to provide a new complementary approach that is capable of reconstructing and analysing such inconspicuous traces that are susceptible to equifinality.

The HIROX is a powerful tool for visualisation purposes, providing high-resolution photographic and observational data which may aid in studies regarding use wear patterns (Stemp et al., 2015; Fernández-Marchena et al., 2016; Martín-Viveros, 2016; Marciani et al., 2018; Martin-Viveros et al., 2018; Martin-Viveros and Ollé, 2018; Wierer et al., 2018), residues (Revedin et al., 2010; Ronchitelli et al., 2015), dental micro-wear analysis (Oxilia et al., 2015), taphonomy (Crezzini et al., 2015; Malassé et al., 2016; Pérez et al., 2017; Rufa et al., 2017; Duches et al., 2018; Moclán et al., 2018; Pineda et al., Under Revision) as well as other sub-disciplines within the archaeological field. Developments in taphonomic morphological analysis using high-tech digital reconstruction methods have made a significant change from the initial approximations to cut mark morphology proposed by Walker and Long (1977). While the HIROX KH-8700 3D Digital Microscope has been used before in similar studies (Blasco et al., 2016; Malassé et al., 2016; Maté-González et al., 2017b, c; Duches et al., 2018), these papers have been limited to either biometric, 2D or even mostly qualitative results. Further issues with this microscope can be seen in its' dependency on a light source to provide accurate 3D reconstructions. In this study we confront these topics, demonstrating the capacity of the 3D digital microscopes, such as the HIROX KH-8700, in processing superficial marks and further highlighting the first morphological characterisation of trampling marks.

2. Materials and methods

In order to compare the quality and efficiency of the HIROX KH-8700 3D Digital Microscope in comparison to the DAVID Structured-Light SLS-2 Scanner, an experiment was devised to see how the positioning of light can affect the morphology of the digitally reconstructed cut mark under observation. Both the 2D 7-landmark cross section model described by Maté-González et al. (2015) and the 3D 13-landmark study as described by Courtenay et al. (2017) were tested in the preliminary phase of this study. Once the most accurate lighting position had been established, the second phase of our experimentation confronted the morphological classification of trampling marks using the HIROX KH-8700 3D Digital Microscope.

2.1. HIROX KH-8700 3D digital microscope

The microscope used in this study is the HIROX KH-8700 3D Digital Microscope (Fig. 1), located at the *Institute of Human Paleoecology and Social Evolution* (IPHES) of Tarragona. This microscope is equipped with a MXG-5000REZ triple objective revolving lens (Fig. 1b), with a magnification range from 35x to 5000x plus a field of view from 8 mm to 0.06 mm at an operable distance of 3.5 mm–10.0 mm (Table 1). The microscope is accompanied by a high intensity LED light source that can be positioned around the object (Fig. 1c and d). This light source provides a temperature of 5700 k, closely portraying daylight colour and producing the highest quality real-time images with no warm up time needed. The HIROX microscope provides the possibility of combining both ring and coaxial light (Fig. 1c), while presenting the possibility of using polarised filters. The built in compact CCD camera projects these images onto a high definition LCD 21.5" monitor with high intensity pixel reproduction as well as the capacity to display up to 16.77 million colours with a contrast ratio of 1000:1 and brightness of 300 cd/m^2. The combination of state of the art hardware and the Genex Engine graphics processor ensures maximum quality when carrying out any type of microscopic analysis. The HIROX is capable of quickly producing 3D digital reconstructions using a combination of quick auto focus and depth synthesis functions. The additional use of the HIROX's tiling function is used to create a mosaic and complete digital reconstruction of the subject under analysis.

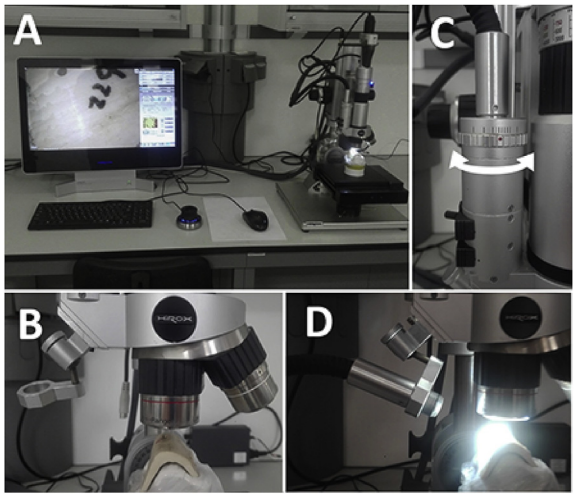


Fig. 1. (A) The HIROX KH-8700 3D Digital Microscope located at the IPHES lab (Tarragona, Spain). (B) The MXG-5000REZ triple objective revolving lens. (C) Lighting position from above, providing the option of combining ring (turning the wheel to the right) and coaxial (turning the wheel to the left) lighting conditions. (D) Lighting position from the side, movable using the HIROX's adjustable light support.

Table 1
Details regarding each of the lenses of the MXG-5000REZ triple objective revolving lens.

Lens Magnification	HFV (Horizontal Field Of View (FOV))	Depth of field	Working distance
35x - 250x	8.76–1.22 mm	0.72–0.072 mm	10 mm
140x - 1000x	2.18–0.31 mm	0.09–0.007 mm	10 mm
700x - 5000x	0.44–0.06 mm	0.01–0.0007 mm	3.5 mm

2.2. DAVID Structured-Light SLS-2 scanner

The DAVID Structured-Light SLS-2 Scanner, located at the TIDOP research group in the University of Salamanca, is a powerful tool in providing real reproductions of external bone topography in less than 1 min. The equipment consists of a DAVID USB CMOS Monochrome camera, an ACER K132 projector and a calibration marker board. This equipment is able to produce a density of up to 1.2 million points while providing a high-resolution 3D model that can be later imported into different graphics software such as Avizo (Visualisation Sciences Group, USA). Use of this equipment in 3D reconstructions of taphonomic traces have already proven to be successful and can be consulted in Courtenay et al. (2017, 2018), Maté-González et al. (2017c) and Yravedra et al. (2018b). Tests regarding superficial taphonomic traces, however, have proven to be unsuccessful (Courtenay et al., 2017).

2.3. Testing the efficiency of the HIROX digital microscope

A total of 5 lighting positions were initially considered (Fig. 1c and d); from the left of the mark, from the right of the mark, a fixed position above the mark using a ring light, a fixed position above the mark using a combination of ring and coaxial light (mix) and finally the use of two light sources from either side of the mark. The use of two lights intended to reduce the amount of shadow cast over the cut mark in an attempt to provide sufficient illumination of the entire incision. The sheer strength of the HIROX Digital Microscope's own light source, however, might present some problems. Multiple types of secondary lights were considered and tested, however, none of these solutions were able to match the intensity of the LED light provided by this microscope. Because of this, the 5th lighting position was discarded.

In order to reduce variability produced by other factors, once the bone was positioned prior to analysis, neither the bone nor the platform could be moved or altered, ensuring that the only changing variable within our analysis was the positioning of the light. Cut mark cross-section profiles were captured using the mid-range lens at a 600× magnification. Between 110 and 130 photos were taken for each profile, combined and constructed into a 3D model using the HIROX's quick auto focus and depth synthesis functions.

For the digital reconstruction of the entire mark, the necessary magnification, as seen by the analyst, was selected in order to capture every feature of the mark, depending on the incisions' depth and size. Each cut mark was digitally reconstructed using the Mosaic 3D Tiling function within the HIROX. The number of photos taken per tile for each mark was used as suggested by the microscope's own internal software, however for higher quality reconstructions a minimum of at least 30 photos is recommended. The total number of tiles was entirely dependent on the microscope and cannot be altered manually by the analyst.

A random selection of cut marks were reconstructed using all 4 different lighting positions. These cut marks, 4 in total, came from previously studied material by Courtenay et al. (2017) and were reproduced using simple quartzite flakes on a *suid* femur diaphysis. These were later processed following the methodological approaches described by both Maté-González et al. (2015) and Courtenay et al. (2017) in order to see how the lighting position affected both the morphology of cut mark section profiles as well as the entire morphology of each mark. Through this, both methods could be assessed, observing the variation in quality of the 3D digital reconstructions of each incision. Additionally, the reliability of the HIROX and its use for geometric morphometrics could be tested.

Notes on the amount of photos taken per tile, the number of tiles used to reconstruct the full incision, the time taken to reconstruct the mark and the time taken for the analyst to process the mark were taken. The number of pixels and the distance between the lowermost and uppermost positioning of the lens were also noted.

2.4. Collecting of landmark data

Cut mark profiles were exported into the free tpsDig2 (v.2.1.7) software where the allocation of 7 homologous landmarks were carried out following the geometric morphometric model described by Maté-González et al. (2015, Fig. 2a).

In order to collect the location of the 13 landmarks according to Courtenay et al. (2017)'s 3D model (Fig. 2b), the position of each point was recorded through a series of measurements. Each measuring tool is provided by the HIROX's own internal software, thus providing accurate results with little to no human error. The self-calibration select sensor within the HIROX's system automatically configures and applies the appropriate lens settings according to the lens and magnification being used, eliminating the need for further calibration. Measurements can be taken using various different tools and simple mouse operations via the monitor's display. The accuracy of these measurements are as small as 1μ.

Landmark coordinate data was first collected using the HIROX's 'XY-Width' function, measuring and plotting across a 2D graph the location of each point (Fig. 3a), followed by a measurement in depth using the 'point height' function (Fig. 3b). In doing so, the landmark's position was established along the *z*-axis of a 3D plot. These landmark coordinates were recorded in a database and later imported into R for statistical analysis.

2.5. Statistics

All statistics were carried out in the free software R (www-rproject.org, Core-Team, 2018).

In order to assess the efficiency of the KH-8700 3D Digital

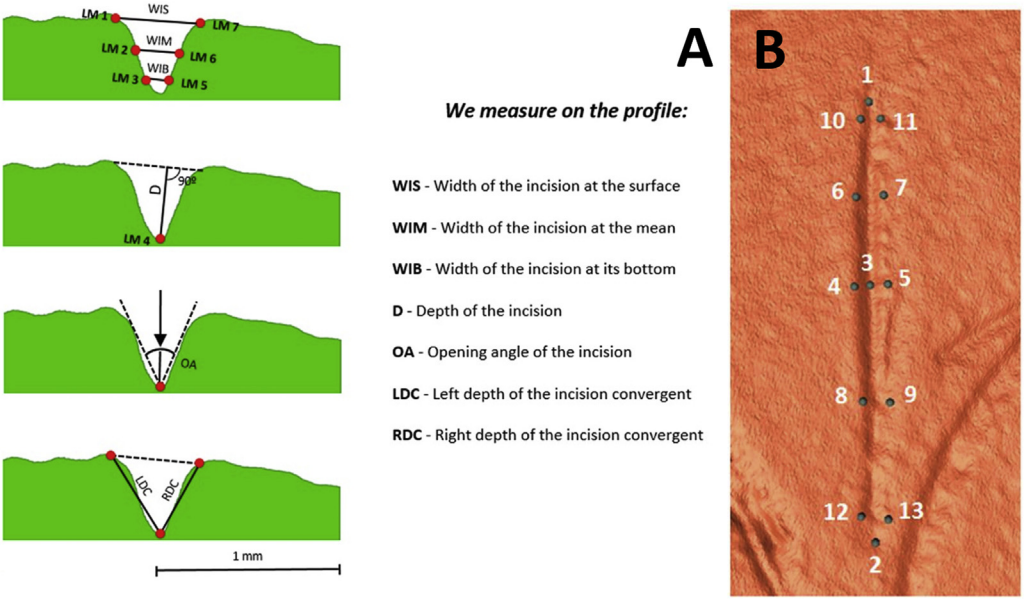


Fig. 2. (A) Location of the seven landmarks used in the 2D morphometric analysis of cut mark cross sections (Maté-González et al., 2015) and the measurements taken for each cut mark profile (Bello and Soligo, 2008). (B) Location of the 13 landmarks used to capture the shape of entire incisions, as described by Courtenay et al. (2017).

Microscope, various polynomial multiple regressions were performed to assess the weight and importance of each variable in the time taken to create a full 3D digital reconstruction of each mark. A step regression was also used to assess which variables were conditioning these results and which were not.

Comparisons of the different lighting conditions on cut mark profiles were firstly performed measuring the thickness, depth and angles of each incision (Fig. 2) following the quantitative methodological approach proposed by Bello and Soligo (2008). This biometric data was obtained using the free software tpsDig2 (v.2.1.7) and later imported into R. This was done in order to test the differences presented in the 3D reconstructions of the cut mark morphology using the different techniques. These measurements were firstly tested for normality according to the Shapiro-Wilk normality test before being subjected to further multivariate analysis (MANOVA). Following this, a jack-knifed linear discriminant analysis was used to determine whether significant differences were present through the calculation of a confusion matrix. For these tests, both the MASS (Venables and Ripley, 2002) and the RVAideMemoire (Hervé, 2018) R packages were used.

Considering previous observations (Maté-González et al., 2017b; Courtenay et al., 2018) on the variables suggested by Bello and Soligo (2008) had a major impact on the variance among groups, these tests were also performed including or excluding the opening angle of the incision (OA).

Geometric morphometric analyses of the cut mark profiles were based on 7 homologous landmarks, as specified above. Files containing

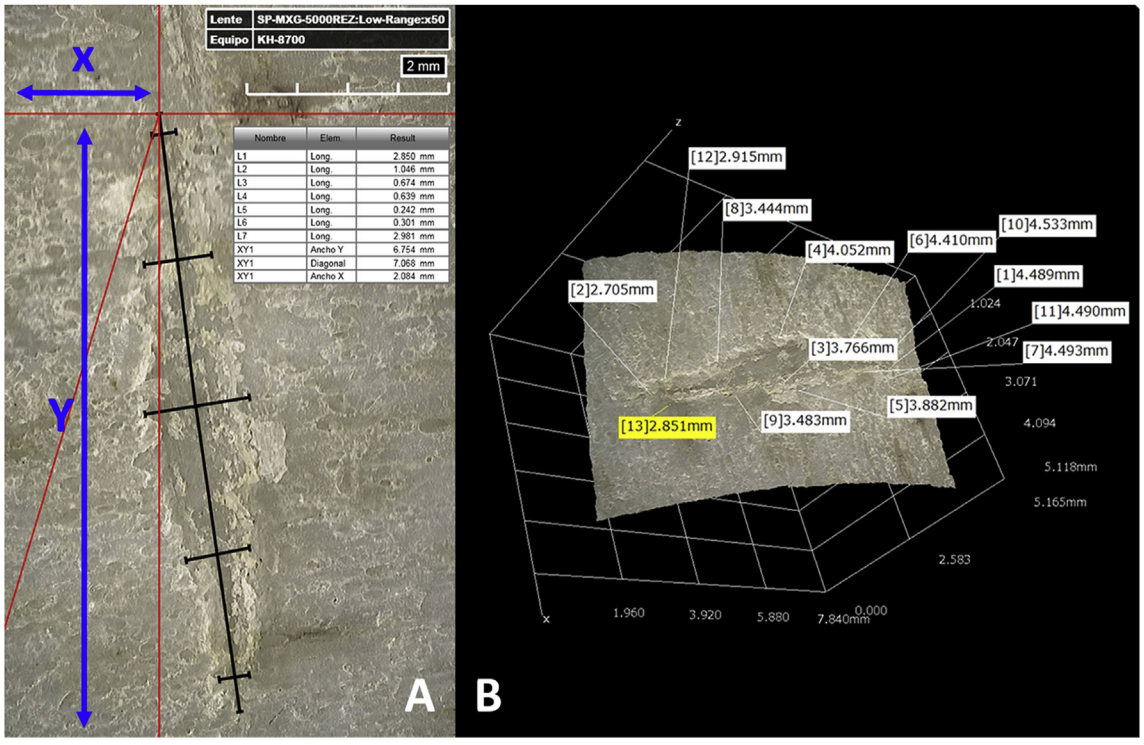


Fig. 3. Description of the measurements taken in order to plot the 3D coordinates of each of the 13 landmarks across the (A) x, y and (B) z axes.

Table 2
–Table presenting the different variables taken when carrying out a 3 Dimensional Reconstruction using the HIROX KH-8700 Digital Microscope and the time taken to process each cut mark under the different lighting conditions.

Cut Mark Number	Lighting Position	Length of Incision (mm)	Width of Incision (mm)	Magnification	Focus Range (Lens Height) (mm)		Number of Pixels		Number of Photos Taken per Tile	Number of Tiles	Tiling Time (Min)	Time Taken to Place Landmarks (Min)	Total Processing Time (Min)
					Sup.	Inf.	X	Y					
1	Ring/ Coaxial	8.494	0.935	200	815.95	−1583.05	2411	9844	25	42	20:47	06:22	27:09
	Coaxial	8.494	0.935	200	815.95	−1583.05	2411	9844	25	42	20:50	04:59	25:49
	Left	8.494	0.935	200	815.94	−1583.05	2411	9844	25	39	19:22	04:59	24:21
	Right	8.494	0.935	200	815.95	−1583.05	2411	9844	25	39	19:10	03:17	22:27
2	Ring/ Coaxial	11.552	1.455	100	3763.6	−470.7	2496	7175	30	27	13:31	07:52	21:23
	Coaxial	11.552	1.455	100	3763.6	−470.7	2496	7175	30	30	14:55	06:15	21:10
	Left	11.552	1.455	100	3763.6	−470.7	2496	7175	30	30	14:40	06:05	20:45
	Right	11.552	1.455	100	3763.6	−470.7	2496	7175	34	23	12:33	07:17	19:50
3	Ring/ Coaxial	6.521	0.682	200	1020.9	−253.3	1731	7479	25	22	08:07	05:20	13:27
	Coaxial	6.521	0.682	200	1020.9	−253.3	1731	7479	25	23	05:32	07:18	12:50
	Left	6.521	0.682	200	1020.9	−253.3	1731	7479	25	23	08:39	05:09	13:12
	Right	6.521	0.682	200	1020.9	−253.3	1731	7479	25	30	08:04	05:30	13:34
4	Ring/ Coaxial	9.000	1.197	150	1531.2	−841.8	2807	7960	25	34	13:24	07:12	20:36
	Coaxial	9.000	1.197	150	1531.2	−841.8	2807	7960	25	34	13:50	06:54	20:44
	Left	9.000	1.197	150	1531.2	−841.8	2807	7960	25	34	14:33	07:48	21:21
	Right	9.000	1.197	150	1531.2	−841.8	2807	7960	25	31	13:23	05:24	18:47

the landmark data were edited and imported into R where a full Procrustes fit and an orthogonal tangent projection (Dryden and Mardia, 1998) were used to normalise the data for further multivariate statistical analyses. This technique, commonly referred to as Generalized Procrustes Analysis (GPA) is used to standardize the form information through the application of superimposition procedures including translation, rotation and scaling. Any remaining differences are exposed through patterns of variation and covariation that can be assessed through several statistical tests (Slice, 2001; Rohlf, 1999).

A principal component analysis (PCA) in shape space and form space on the Procrustes superimposed landmarks was performed using the geomorph package in R (Adams et al., 2017). Changes in both shape and form were visualized with the aid of transformation grids (Bookstein, 1989). PC scores were then extracted and used to examine the variance between samples by means of a MANOVA test and a jack-knifed LDA. A canonical variance analysis (CVA) using 1000 permutations tests was used to determine further morphometric differentiations between the techniques.

3D landmarks were statistically analysed in a similar manner to the 2D landmarks model. This 3D landmark configuration consists of 13 identical points on the exterior and interior surface of each mark, following Courtenay et al. (2017)'s methodological approach. Geometric morphometric analysis began, as before, with a GPA based on the natural logarithm of centroid size. PCAs were later produced in order to assess patterns of variation, with the help of transformation grids and warpings, thus revealing any differences that may be produced through the change in lighting position. PC scores were later extracted and used to assess the similarities and differences between methods based on MANOVA tests, LDA jack-knifed confusion matrices and finally CVA tests, as described before.

Once the best lighting position was established, all subsequent digital reconstructions followed this exact method.

2.6. Trampling experiments

Trampling experiments were carried out using sieved sediments from levels 2 and 3 from the archaeological site of Peña de Estebanvela (Segovia, Spain) (Yravedra, 2005, pp. 249–253). This site has been dated using calibrated AMS ^{14}C to around 10–14 Ka (Jordá Pardo and Cacho, 2013), consequently being attributed to various phases of the Magdalenian (Cacho et al., 2016; Yravedra et al., 2018). Sediment samples obtained from this site consist of a mixture of compact and loose quartz sandy sediments that were used to bury the remains of an *Ovis* skeleton. These sandy sediments are composed of quartz granules, presenting an angular morphology with an average granular size of 250 μm. In total 14 bones of both axial and appendicular skeletal elements were buried. These consisted of mandibulae (Number of Elements (NE) = 2), femorae (NE = 2), tibiae (NE = 2), and radiae (NE = 2). All the bones were dry and meatless when buried while the sediments used to bury the bones were also dry. Once the bones were buried, the area was trampled on by a single individual for a total of 5 min before being uncovered, cleaned and studied.

A total of 56 trampling marks were produced, however, not all presented clear morphologies where all 13 landmarks could be easily located and processed. For the purpose of this study, and under the premise of obtaining a statistically significant sample for morphological characterisation, 30 trampling marks were carefully selected and processed.

Due to their superficial nature, these trampling marks present a practically non-existent cross section to study. As a product of this, the 3D 13-landmark model proposed by Courtenay et al. (2017) was solely used to process each trampling mark. The landmark data was then statistically processed in the free software R, as previously described.

3. Results

3.1. Comparing the David SLS-2 versus the HIROX KH-8700

3.1.1. Time

The HIROX was able to recreate each incision in a time range of 5:32 to 20:50 min with an average time of 13:30 min (Table 2). Compared with the processing time required by the DAVID SLS-2 (less than a minute) (Courtenay et al., 2017; Maté-González et al., 2017c), the HIROX is substantially slower (Table 3).

The use and comparison of multiple linear, logistical and polynomial regressions have shown that the most important conditioning variables in the time taken to reconstruct each mark are, in order of importance; cut mark length, magnification, the number of tiles and finally the cut mark's width (AIC = −1.62). The use of stepwise regressions and consequent comparison of Akaike Information Criterion

Table 3
Comparisons between different reconstruction techniques. Updated from Maté-González et al. (2017c).

Technique	System	Measuring Procedure	Classification	Portability	Full 3D Reconstruction Time (Min)	Operable Distance	Resolution	Cost (Eur)
Microscope	KH-8700	3D Digital Microscope	Active Sensor	Low	Aprox. 13.5	1–10 mm	0.15–0.01 μm	< 100,000
Laser Scanner	David SLS-2	Structured Light	Active Sensor	Medium	> 1	0.15–5.00 m	0.02 mm	3000
Photogrammetry	Reflex + Macro Objective	Micro-photogrammetry	Passive Sensor	High	Aprox. 25	10–50 cm	0.02 mm	1000

(AIC) values confirm that these 4 variables have a significant ($p < 0.05$) effect on the processing time to reconstruct the cut marks. Since the number of tiles is entirely dependent on the magnification, the length and width of each incision, a clear correlation between these variables and the reconstruction conditions of the taphonomic traces can be observed.

In contrast, the time required for the landmarking process does not vary substantially between the two methods (HIROX vs DAVID SLS-2). Therefore, the time taken to landmark the traces has not been accounted for in this study.

3.1.2. Lighting positions

Measurements taken from cut mark profile reconstructions present limited variability. Results obtained through MANOVA ($p \approx 1$) and LDA tests support this statement (Table S1). This proves that the reconstructions generated using different lighting positions with the HIROX are practically indistinguishable from those obtained using the laser scanner. The exclusion the OA measurement provides a slightly higher level of classification (15%) in the LDA (Table S1), however MANOVA results still highlight the similarities of both reconstruction methods ($p \approx 0.9$).

2D geometric morphometric analysis produced a total of 10 PC scores (Fig. 4a), and a 2D plot where most groups overlap with a wide dispersion of individuals across tangent space regardless of the reconstruction technique used. In general, these tests do not show a clear differentiation between each cut mark, making it hard to distinguish between the different profiles. MANOVA results present insignificant differences between the lighting positions (Table 4) yet are also inconclusive in distinguishing between the different cut marks regardless of the technique used for the reconstruction (Table 5). Considering distance in shape space across the PCA graph, lighting position from above (both mix and ring) present the highest level of similarities with the laser scanner. The same can be said considering MANOVA results.

CVA results represent a large dispersion of points (Fig. 4b), where the differentiation between groups appears to be slightly clearer; probably because CVA tests tend to overestimate differences. LDA classification/misclassification tables present a relatively high percentage of correctly classified cut marks (Table S4), while the reconstruction techniques in general are considered highly similar to those of the scanner (Table S5).

The PCA on the 3D reconstructions is defined by a total of 19 PC scores. The scatter-plot containing the first two PCs shows that cut marks can be clearly differentiated regardless of the lighting position (Fig. 5a). These results are strongly supported by MANOVA, which present high similarities amongst all 5 reconstruction techniques ($p > 0.4$) and are still able to distinguish between the cut marks in most cases ($p < 0.05$). The only exceptions are found in two of the cut marks that can be considered practically identical in morphology. CVA results present 3CV scores (CV1 and CV2 = 94% cumulative variance) that clearly differentiate between all cut marks (Fig. 5b), tightly grouping all 4 reconstruction techniques. LDA jack-knifed classification/misclassification results highlight that all 4 cut marks are completely distinguishable, indicating that despite the use of the HIROX or the David SLS-2, the 13-landmark 3D model is capable of distinguishing perfectly between cut marks (Table S4). LDA results comparing reconstruction techniques present high percentages of misclassification in all cases, highlighting a lack of differentiation between techniques and thus implying the similarities in reconstruction quality for all cases (Table S5).

2D variations in shape (Fig. 4) are mostly defined by differences in cut mark depth (PC1), and in the angle of the mark and irregularities on the incision wall (PC2). 3D transformation grids along PC1 indicate changes in the curve and depth of the cut mark, whereas PC2 underlines a slight variation in the positioning of landmark 3 (indicating a slight change in angle) and in the width of the cut mark. Considering results presented by transformation grids, changes in directionality tend to cast a shadow over the profile, thus affecting the angle of the mark. While this can be problematic for 2D reconstructions, the effects on 3D landmark models is minute and reflected through one single landmark. These results are strongly supported by both significant numerical and graphic results, concluding that the most accurate reconstruction technique employed the use of a lighting position from above. The mixture of coaxial and ring lighting were considered optimum for higher magnifications, such as those required for the study of superficial taphonomic traces. Thus, this technique was preferred for the rest of this study.

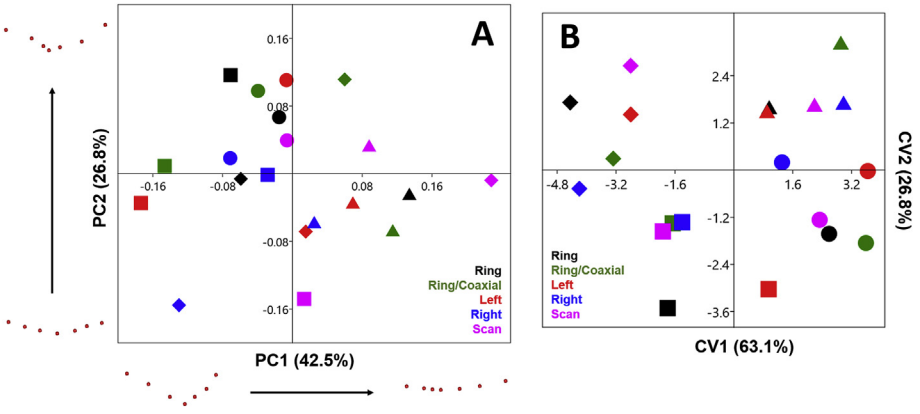


Fig. 4. Two scatter plots presenting (A) PCA and (B) CVA graphs comparing the 5 different reconstruction techniques of different cut mark cross sections according to the 2D 7-landmark model as described by Maté-González et al. (2015). Each cut mark is represented by a different symbol while variance in shape is presented for the extremities of both PC scores along their respected axis.

Table 4
MANOVA p values comparing the different reconstruction techniques, both using the 2D landmark model of cut mark profiles (in non-bold typeface) and using the 13 landmark 3D model of the entire incision (in bold typeface).

	Ring	Mix	Left	Right	Scan
Ring		0.98689 **0.94135**	0.86246 **0.77183**	0.70995 **0.90411**	0.85949 **0.40608**
Mix	0.98689 **0.94135**		0.72821 **0.98715**	0.62274 **0.99989**	0.83721 **0.61758**
Left	0.86246 **0.77183**	0.72821 **0.98715**		0.9868 **0.99242**	0.57261 **0.74994**
Right	0.70995 **0.90411**	0.62274 **0.99989**	0.9868 **0.99242**		0.47208 **0.64374**
Scan	0.85949 **0.40608**	0.83721 **0.61758**	0.57261 **0.74994**	0.47208 **0.64374**	

3.2. Digital reconstruction protocol

Through the statistical results presented here and experience using the HIROX, the ideal conditions for 3D digital reconstruction are as follows:

After having calibrated and centralised the main table, the piece should be placed on a sturdy support, specimen mount or directly on the table with the area of interest positioned as flat as possible. This ensures that in the process of reconstruction the piece is unable to move, thus preventing blurry photos. The cut mark should be positioned either vertically or longitudinally as straight as possible, therefore reducing the number of tiles needed in the digital reconstruction. Additionally, the reconstruction should require the least amount of focal depth possible. In this experiment, the focal range oscillated between 1 and 3 mm (Table 2), however, superficial marks that require high levels of magnification ideally require the lowest focal depth possible. This particular variable, however, is highly dependent on the natural topography of the piece, its' positioning on the table and can also vary due to the magnification used. The analyst is therefore advised to ensure that the piece is as flat as possible, while consulting the depth of focus of the lens used, thus adjusting accordingly (Table 2).

In order to capture the entire morphology of the object under study, sufficient magnification is essential in order to clearly observe, photograph and thus capture the base, walls and surrounding cortical of each incision. For 2D reconstruction of cross sections, 600× magnification (FOV = 505 µm, Table 1) using the mid-range lens is ideal. The section profile produced should be captured between 30% and 50% of the entire mark's length, as initially described by Maté-González et al. (2015). In this case the mosaic function is not necessary. For 3D reconstructions of entire marks, under most circumstances the ideal magnification can be considered between 100x (FOV = 1516 µm, Table 1) and 200x (FOV = 3032 µm, Table 1) magnification. This can be performed using either the low-range or the mid-range lens. The entire mark can be captured using the mosaic function, however the analyst has no control over the number of tiles necessary to fully reconstruct the mark. Depending on the magnification, number of tiles and the dimensions of the object under study, the time required to create the 3D reconstruction will vary.

The number of photographs taken per tile is strongly recommended to be as few as 30, thus ensuring the best perception of depth for each mark. The optimum lighting condition is with the light source placed directly above the sample, using a mixture of coaxial and ring light according to the magnification used. No polarising filters are advised at magnifications lower than 1000x.

3.3. Analysis of trampling

The HIROX KH-8700 performed efficiently enough to reconstruct each trampling mark (n = 30) and to collect the required 3D landmark data. During the analysis of trampling marks, a series of patterns were observed, possibly highlighting the presence of two different groups of trampling marks in the sample. Both groups can be qualitatively characterised by their width and the quantity of internal striae either on the floor or along the wall of the mark. The first group consists of thinner traces with very few or no internal striations (Fig. 6a), whereas the marks included in the second group are wider and present a high number of internal striations (Fig. 6b). The qualitative identification of these two groups was further tested through the use of a K-Means clustering model (Table 6). This test was carried out on the PC scores obtained through PCA which, in turn, confirmed the presence of two

Table 5
MANOVA p values comparing the differentiation of different cut marks using varied reconstruction techniques, both using the 2D landmark model of cut mark profiles (in non-bold typeface) and using the 13 landmark 3D model of the entire incision (in bold typeface).

	Cut Mark 1	Cut Mark 2	Cut Mark 3	Cut Mark 4
Cut Mark 1		0.64485 **0.007125**	0.15693 **0.009103**	0.14058 **0.001474**
Cut Mark 2	0.64485 **0.0071251**		0.15665 **0.078724**	0.15431 **0.002631**
Cut Mark 3	0.15693 **0.0091033**	0.15665 **0.078724**		0.40692 **0.00753**
Cut Mark 4	0.14058 **0.0014735**	0.15431 **0.002631**	0.40692 **0.00753**	

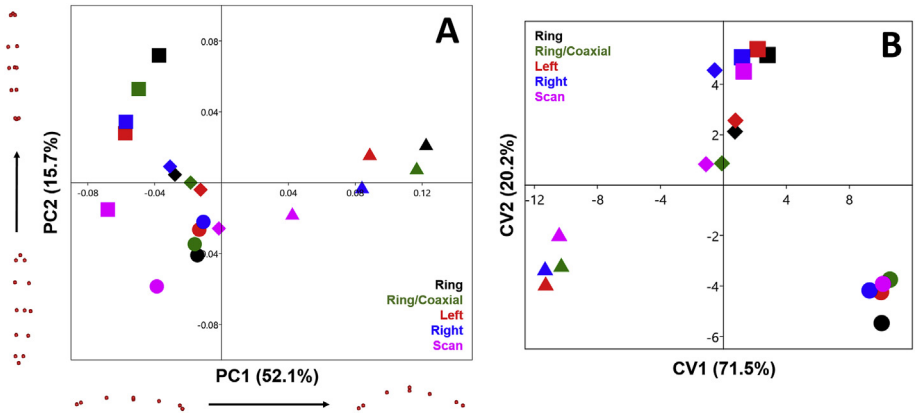


Fig. 5. Two scatter plots presenting (A) PCA and (B) CVA graphs comparing the 5 different reconstruction techniques of different cut marks according to the 3D 13-landmark model as described by Courtenay et al. (2017). Each cut mark is represented by a different symbol while variance in shape is presented for the extremities of both PC scores along their respected axis.

Fig. 6. Examples of the two different types trampling marks: (A) scratches and (B) grazes.

different groups of trampling marks based on their morphology.

The results for the K-Means clustering model were then applied to future morphological analysis to further classify these two different trampling groups. In total the clustering model separated 21 of the 30 trampling marks into one group and the remaining 9 were separated into the second group. The first group, consisting of the much finer trampling marks, were named '*scratches*' (Fig. 6a), whereas the wider marks with abundant internal striae were named '*grazes*' (Fig. 6b). The term *scratch* was assigned based on the etymological definition of the word as a "score or mark [of] the surface of (something) with a sharp pointed object" (The Oxford English Dictionary, 2018). The term *graze* was selected as a result of its definition as a "scrape or break [of] the surface of [something]" or a "touch or scrape lightly in passing" (The Oxford English Dictionary, 2018), thus deemed appropriate, considering the connotations associated with the word *graze*.

PCA results and transformation grids (Fig. 7) confirm the original hypothesis that these marks can be morphologically characterised and separated through their width. Additionally, the variance in shape of the mark's trajectory is reflected across PC2 as an important variable to define trampling marks. This confirms Domínguez-Rodrigo et al. (2009)'s original characterisation of trampling marks through their either sinuous or curvy groove trajectory.

CVA results clearly separate the two trampling groups based on their shape through significant Mahalanobis ($D = 19.7491$, $p < 0.0001$) and Procrustes ($D = 0.0570$, $p < 0.0001$) distances. This separation is clearly supported by MANOVA results ($p = 9.12e\text{-}06$) as well as jack-knifed LDA classification/misclassification tables with a correct classification range between 93.33% of the samples (Table 6).

Differences in form space do not vary much from results obtained in shape space (Fig. 8), except for a slight overlapping between groups. In this case, transformation grids accentuate a change in the width of the trampling marks across both PC1 and PC2. A slight difference can also be seen in the distance between landmarks 10/11 and 12/13 with the corresponding end of each groove, indicating that the length of *scratches* is longer than those of *grazes*. MANOVA results still identify significant differences between groups ($p = 1.0543e\text{-}05$) while jack-knifed LDA classification tables correctly classify 76.67% of the sample (Table 6).

Table 6
LDA Classification/Misclassification table presenting the possibility of correctly associating each mark to the respected K-Means cluster group.

	Shape			Form		
K-Means	LDA	Graze	Scratch	LDA	Graze	Scratch
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	31%	69%	Graze	66%	34%
Scratch	Scratch	2%	98%	Scratch	27%	73%
Graze	Graze	83%	17%	Scratch	83%	17%
Scratch	Scratch	0%	100%	Graze	9%	91%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Graze	Graze	100%	0%	Scratch	20%	80%
Graze	Scratch	47%	53%	Graze	86%	14%
Scratch	Scratch	5%	95%	Scratch	1%	99%
Graze	Graze	97%	3%	Scratch	14%	86%
Scratch	Scratch	22%	78%	Graze	100%	0%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Graze	Graze	100%	0%	Graze	100%	0%
Scratch	Scratch	0%	100%	Scratch	2%	98%
Graze	Graze	69%	31%	Graze	78%	22%
Scratch	Scratch	10%	90%	Scratch	7%	93%
Scratch	Scratch	2%	98%	Scratch	5%	95%
Scratch	Scratch	1%	99%	Scratch	5%	95%
Scratch	Scratch	1%	99%	Scratch	1%	99%
Scratch	Scratch	0%	100%	Scratch	6%	94%
Graze	Graze	99%	1%	Graze	94%	6%
Graze	Graze	98%	2%	Graze	87%	13%
Graze	Scratch	1%	99%	Scratch	29%	71%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	1%	99%	Scratch	10%	90%
Scratch	Scratch	0%	100%	Scratch	1%	99%
Scratch	Scratch	5%	95%	Scratch	11%	89%
Scratch	Scratch	0%	100%	Scratch	0%	100%

4. Discussion

Microscopy in archaeology is an important tool when considering the amount of information available through minute BSMs that remain invisible to the naked eye. This study presents the potential that the 3D digital microscopes have to offer, providing a solution to the problems in resolution presented by laser scanners such as the DAVID SLS-2 when studying superficial taphonomic traces. Through this we have been able

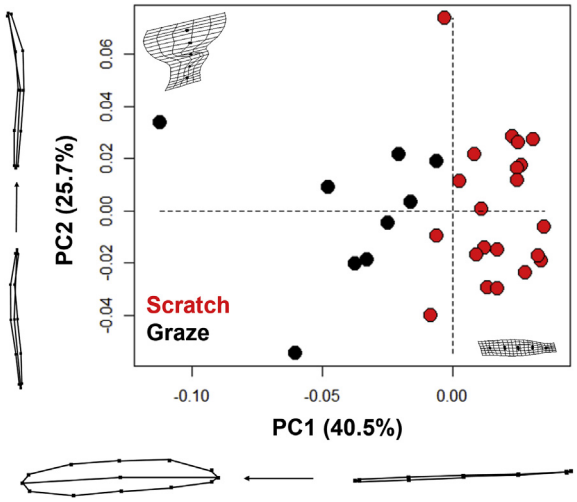


Fig. 7. Scatter plot of the PCA results comparing the morphology of both scratches and grazes in shape space. Variance in shape is presented for the extremities of both PC scores along their respected axis.

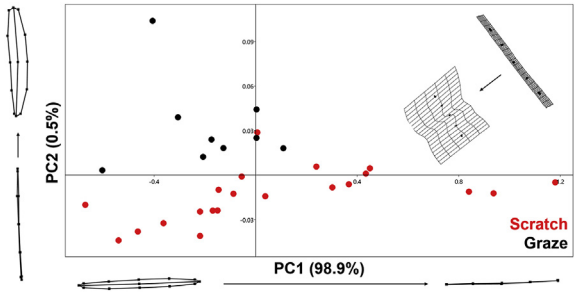


Fig. 8. Scatter plot of the PCA results comparing the morphology of both scratches and grazes in form space. Variance in form is presented for the extremities of both PC scores along their respected axis.

to characterise trampling marks through their morphology, revealing two new categories of naturally produced marks, hereby named *scratch* and *graze* trampling marks.

The impact the two new trampling marks may have on taphonomy is still to be tested. Traditionally, it would be assumed that the nature of trampling marks are conditioned by sedimentological and geological features (Schiffer and Shipman, 1987; Olsen and Shipman, 1988; Fisher, 1995; Marín-Monfort et al., 2013; Reynard, 2014). This may be the explanation behind the occurrence of two different types of morphologically comparable trampling marks. Depending on the abrasive nature of the sediment, different frequencies of *scratch* and *graze* marks could be produced. In our experiment, a mixture of different fine sands have been used, however, gravel like sediments and layers of large quartz granules are more likely to cause more damage to bone cortical surfaces (Marín-Monfort et al., 2013; Reynard, 2014). Additionally, other factors such as bone density, cortical hardness and time exposed to trampling processes may also need to be considered (Öhman et al., 2012; Walden et al., 2017). These variables, however, require a wide range of experimental studies in order to empirically answer these questions.

Methodologically, while the HIROX KH-8700 3D Digital Microscope provides some limitations (Maté-González et al., 2017a, b), considering the price and analytical speed of this equipment, our study highlights a series of advantages. Information revealed through statistics as well as the comparison of transformation grids and warpings highlight that regardless of the lighting position, the HIROX does not provide too much variation in digital reconstructions. While 3D reconstructions appear to be the least affected by the difference in light source, 2D analysis of cut mark cross sections are more susceptible to change when positioning the HIROX's LED light at either side of the mark. This might be explained by the shadows cast across the section of each mark. Statistically, the changes in reconstruction techniques are mostly insignificant. However, in order to present a more accurate representation of mark morphology, the lighting position must remain homogenous throughout any study. With regards to an ideal lighting condition we recommend using a lighting position from above, using a combination of ring and coaxial light in order to properly illuminate the entire mark.

Furthermore, as can be seen through these results, the HIROX digital microscope is much more powerful at analysing more inconspicuous marks than the DAVID SLS-2. A clear disadvantage of this approach, however, can be seen in the amount of processing time required. An approximately 2 s procedure with the laser scanner can be considered much more efficient when analysing large samples, however, when regarding problems related to equifinality, time should not be favoured over resolution. A further advantage of the DAVID SLS-2 over the HIROX is the cost (Table 3). Nevertheless, the visual advantages presented by the HIROX microscope are much greater, especially when considering the perception of texture and the versatility for further qualitative analysis (Fig. 9).

In addition to this, the HIROX's own mosaic tiling functions provide a much more efficient means of capturing larger surface areas, as opposed to microscopes such as the SEM (Vergès and Morales, 2014), or even photogrammetry (Maté-González et al., 2015; González-Aguilera et al., 2016), that rely on longer additional steps in order to fully reconstruct the area of study. The recent development of SEM 3D Images (Eulitz and Reiss, 2015; Tafti et al., 2015), however, present a powerful advance for microscopy. The use of microscopes such as the HIROX to extract linear mark cross sections from archaeological remains has increased over the past years, generating qualitative conclusions drawn from arguably objective means of obtaining this information. Work by Blasco et al. (2016) combine qualitative criteria (Domínguez-Rodrigo et al., 2009) with observations of a homogeneous \/shaped groove to argue their classification of taphonomic traces as cut marks. Rodríguez-Hidalgo et al. (2018) further classify the \/shaped cross section by calculating mean shapes (via landmark data) along the length of the

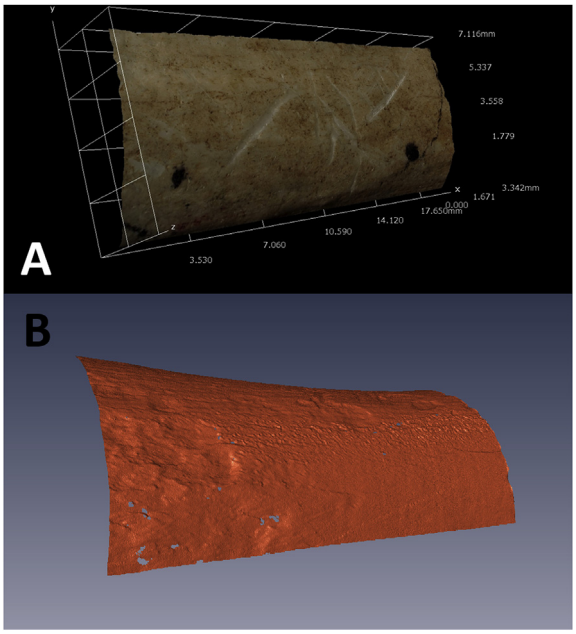


Fig. 9. A comparison of different 3D digital reconstructions using (A) the HIROX KH-8700 3D Digital Microscope and (B) the David Structured-Light SLS-2 Scanner.

groove. This particular approach is strongly supported by quantitative data, yet with the support of an experimental comparative sample has the potential of revealing very important information about the taphonomic register of this site. Similar studies use simpler measurements and calculations to withdraw interesting results (Moretti et al., 2015; Fuentes-Sánchez et al., 2017; Rufa et al., 2017; Duches et al., 2018; Stinnesbeck et al., 2018) that bring taphonomic analysis into the 21st century.

Other notable publications in the field of taphonomy have presented different means of classifying and distinguishing taphonomic traces using quantitative methods, relying on 3D digital reconstructions using confocal profilometers (Pante et al., 2017; Gümrükçu and Pante, 2018; Orlikoff et al., 2018). While data regarding trampling and fluvial altered surfaces provide promising results, comparisons between carnivore tooth marks and cut marks do not respond to any real archaeological questions. Furthermore, distinguishing between these traces can be considered unnecessary when little equifinality is present between the two (Domínguez-Rodrigo and Baquedano, 2018), and are easily macroscopically distinguishable. Observations made by said authors, with the aid of an increased sample size, could begin to respond to important taphonomic issues that have been highlighted in recent years (Pineda et al., 2014; Under Revision).

While microscopy provides a solid methodological groundwork on which to build upon, caution is still advised when relying solely on this data, especially in cases where high-powered technology is used to withdraw rather questionable results (Malassé et al., 2016) based on almost purely qualitative data and inadequate experimental comparative samples. In recent years, some works have diminished the subjectivity present in archaeological and taphonomical studies (Arriaza and Domínguez-Rodrigo, 2016; Egeland et al., 2018; Domínguez and Baquedano, 2018), based on autocritical research presented by Domínguez-Rodrigo et al. (2017, 2018). The level of subjectivity in our work is arguably dependent on an analyst's experience and knowledge, however the use of advanced technology have begun to present ways of overcoming analyst subjectivity. Advances in Artificial Intelligence (AI) and Machine Learning (ML) methods applied to archaeology have begun to break new grounds (Domínguez-Rodrigo, 2018), presenting powerful statistical means of reanalysing data and confronting taphonomic issues (Arriaza and Domínguez-Rodrigo, 2016; Domínguez and Baquedano, 2018), additionally drawing to light interpretation problems that require further attention (Egeland et al., 2018).

5. Conclusions

The 21st century has seen the arrival of technological advances that have revolutionised science. In archaeology, the impact of this paradigm shift can be observed in the application of new technologies to the study of archaeological sites. This paper presents the HIROX KH-8700 3D Digital Microscope as an important tool that is capable of analysing superficial taphonomic traces alongside a new characterisation of naturally produced trampling marks through their morphologic features. The methodological approach described here can be combined to decipher the taphonomic register present within a site from a much more objective perspective. Additionally, applied statistics and the use of landmark data in geometric morphometric analyses are decreasing the degree of equifinality present in taphonomic assemblages.

While the majority of efforts have been focused on the analysis of cut marks, and more recently on carnivore BSMs, few efforts have gone into the geometric morphometric characterisation of naturally produced marks such as trampling. While this study begins to confront these concepts, an ideal progression from this would be to combine these models with our current understanding of cut marks, thus beginning to eliminate a great cause for confusion that many taphonomists face. Differentiation between these two taphonomic traces via geometric morphometrics would be a considerable advance in taphonomic research. Additionally, further investigation into the nature behind *scratch* and *graze* marks may also provide valuable information into effects of sedimentology on trampling marks.

Nevertheless, the impact microscopy has on archaeology is not exclusively limited to taphonomy. Our newfound understanding of this equipment can be applied to other fields of research as well, including traceology, anthropology, and many more disciplines in prehistoric research.

These results present an interesting starting point, especially when employing a combined usage of the 3D digital microscopy with other equipment such as structured light scanners. Considering the results produced in previous research comparing these pieces of equipment, it would be interesting to further include other techniques such as confocal profilometers in this field, to fully understand the advantages of each method. Confocal profilometers are yet to be fully compared with microscopes and, taphonomically, have only been used to test macroscopically distinguishable marks where little equifinality is present. It would be of great interest to investigate the degree of resolution that such equipment can provide, especially if applied to superficial taphonomic traces such as trampling marks.

Acknowledgements

We would like to thank all the staff and members of both the IPHES and the Rovira i Virgili University for helping provide the essential tools we needed in carrying out this study. We would also like to thank Jordan Courtenay for her support and her advice when proofreading sections of this paper.

Additionally, we would like to thank the TIDOP Group from the Department of Cartographic and Land Engineering of the Higher Polytechnics School of Avila, University of Salamanca, for the use of their tools and facilities. We want to recognize the technical support provided by C.A.I. Arqueometry and Archaeological Analysis from Complutense University which has been very useful in carrying out the present paper. J.A. would like to thank Fundación La Caixa and the Spanish Education, Culture and Sports Ministry (FPU15/04585) for funding her postgraduate education program. Finally, this work was developed within the general framework of the Spanish MINECO-FEDER project CGL2015-65387-C3-1-P, the Catalan AGAUR project 2017-SGR-1040, and the URV project 2017-PFR-URV-B2-91.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.quaint.2018.12.019.

References

Adams, D.C., Collyer, M.L., Kaliontzopoulou, A., Sherratt, E., 2017. Geomorph: Software for Geometric Morphometric Analysis. R Package Version 3.0.5. http://cran.r-project.org/package=geomorph.

Albuquerque, Shipman, P., 1981. Life History of a Fossil. An Introduction to Taphonomy and Paleoecology. Harvard, University Press, London.

Andrews, P., Cook, J., 1985. Natural modifications to bones in a temperate setting. Man 20 (4), 675–691.

Aramendi, J., Maté-González, M.A., Yravedra, J., Cruz Ortega, M., Arriaza, M.C., González-Aguilera, D., Baquedano, E., Domínguez-Rodrigo, M., 2017. Discerning carnivore agency through the three-dimensional study of tooth pits: revisiting crocodile feeding leistoc at FLK-Zinj and FLK NN3 (Olduvai Gorge, Tanzania). Palaeogeogr. Palaeoclimatol. Palaeoecol. https://doi.org/10.1016/j.palaeo.2017.05.021.

Archer, W., Braun, D.R., 2013. Investigating the signature of aquatic resource use within pleistocene hominin dietary adaptations. PLoS One 8 (8), e69899. https://doi.org/10.1371/journal.pone.0069899.

Arriaza, M.C., Domínguez-Rodrigo, M., 2016. When felids and hominins ruled at olduvai gorge: a machine learning analysis of the skeletal profiles of the non-anthropogenic bed I sites. Quat. Sci. Rev. 139, 43–52.

Arriaza, M.C., Yravedra, J., Domínguez-Rodrigo, M., Maté-González, M.Á., Vargas, E.G., Palomeque-González, J.P., Aramendi, J., González-Aguilera, D., Baquedano, E., 2017. On applications of micro-photogrammetry and geometric morphometrics to studies of tooth mark morphology: the modern olduvai carnivore site (Tanzania). Palaeogeogr. Palaeoclimatol. Palaeoecol. https://doi.org/10.1016/j.palaeo.2017.01.036.

Behrensmeyer, A.K., 1978. Taphonomic and ecologic information from bone weathering.

Paleobiology 4 (2), 150–162.
Behrensmeyer, A.K., Gordon, K.D., Yanagi, G.T., 1986. Trampling as a cause for bone surface damage and pseudo-cutmarks. Nature 319, 768–771.
Bello, S., 2011. New results from the examination of cut-marks using three-dimensional imaging. In: Ashoton, N.M., Lewis, S.G., Stringer, C.B. (Eds.), The Ancient Human Occupation of Britain. Elsevier, Amsterdam, pp. 249–262.
Bello, S.M., Soligo, C., 2008. A new method for the quantitative analysis of cutmark micromorphology. J. Archaeol. Sci. 35 (6), 1542–1552.
Bello, S., Parfitt, S., Stringer, C., 2009. Quantitative micromorphological analyses of cut marks produced by ancient and modern handaxes. J. Archaeol. Sci. 36 (9), 1869–1880.
Bello, S., Groote, I., Delbarre, G., 2013. Application of 3-dimensional microscopy and micro-CT scanning to the analysis of Magdalenian portable art on bone antler. J. Archaeol. Sci. 40 (5), 2464–2476.
Bello, S.M., Delbarre, G., de Groote, I., Parfitt, S.A., 2016. A newly discovered flint-knapping hammer and the question of their rarity in the palaeolithic archaeological record: reality or bias? Quat. Int. 403, 107–117.
Binford, L., 1967. Smudge pits and hide smoking. The use of analogy in archaeological reasoning. Am. Antiq. 32 (1), 2–12.
Binford, L., 1968. Archaeological Perspectives, New Perspectives in Archaeology. Aldine, New York, pp. 5–32.
Binford, L., 1981. Bones: Ancient Men and Modern Myths. Academic Press Inc, New York.
Blasco, R., Rosell, J., Smith, K.T., Maul, L.C., Sañudo, P., Barkai, R., Gopher, A., 2016. Tortoises as a dietary supplement: a view from the middle pleistocene site of qesem cave, Israel. Quat. Sci. Rev. 133, 165–182.
Blumenschine, R., 1986. Early Hominid Scavenging Opportunities. Implications of Carcass Availability in the Serengeti and Ngorongoro Ecosystems. Archaeopress, Oxford.
Blumenschine, R., 1995. Percussion marks, tooth marks, and experimental determinations of the timing of hominid and carnivore access to Long bones at FLK zinajanthropus, olduvai gorge, Tanzania. J. Hum. Evol. 29 (1), 21–51.
Blumenschine, R., Prassack, K.A., Kreger, D., Pante, M.C., 2007. Carnivore tooth-marks, microbial bioerosion, and the invalidation of domínguez-rodrigo and barba's (2006) test of oldowan hominin scavenging behavior. J. Hum. Evol. 53 (4), 420–426.
Bookstein, F., 1989. Principal warps: thin-plate splin and decomposition of deformations. Transactions on Pattern Analysis and Machine Intelligence 11 (6) 597-585.
Brain, C.K., 1967. Bone weathering and the problem of pseudo-tools. South Afr. J. Sci. 63, 97–99.
Bunge, M., 1981. Analogy between systems. Int. J. Gen. Syst. 7, 221–223.
Cacho, C., Martos, J.A., Yravedra, J., Ortega, P., Martín-Lerma, I., Avezuela, B., Zapata, L., Ruiz-Alonso, M., Valdivia, J., 2016. On the use of space at La Peña de Estebanvela (ayllón, segovia, Spain): an approach to economic and social behaviour in the upper magdalenian. Quat. Int. 412, 44–53.
Cook, J., 1986. Marked human bones from gough's cave, somerset. Proc. - Univ. Bristol Spelaeol. Soc. 17 (3), 275–285.
Core, R.Team, 2018. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.
Courtenay, L.A., Yravedra, J., Maté-González, M.Á., Aramendi, J., González-Aguilera, D., 2017. 3D analysis of cut marks using a new geometric morphometric methodological approach. Journal of Archaeological and Anthropological Sciences. https://doi.org/10.1007/s12520-017-0554-x.
Courtenay, L.A., Maté-González, M.Á., Aramendi, J., Yravedra, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2018. Testing accuracy in 2D and 3D geometric morphometric methods for cut mark identification and classification. PeerJ 6, e5133. https://doi.org/10.7717/peerj.5133.
Crezzini, J., Boschin, F., Arrighi, S., Giamello, M., Ronchitelli, A., 2015. Zooarchaeology, Prehistoric Art and digital imaging: quantitative analysis of bone modifications using 3D microscopy. In: Conference: 43rd Computer Applications and Quantitative Methods in Archaeology.
de Juana, S., Galán, A.B., Domínguez-Rodrigo, M., 2010. Taphonomic identification of cut marks made with lithic handaxes: an experimental study. J. Archaeol. Sci. 37, 1841–1850.
Domíguez-Rodrigo, M., Barba, R., 2007. Five More Arguments to Invalidate the Passive Scavenging Version of the Carnivore-Hominid-Carnivore Model: a Reply to Blumenschine et al. (2007a). J. Hum. Evol. 53 (4), 427–433.
Domínguez-Rodrigo, M., 1997. Meat-eating by early hominids at the FLK 22 zinjanthropus site, olduvai gorge, Tanzania: an experimental approach using cut mark data. J. Hum. Evol. 33 (6), 669–690.
Domínguez-Rodrigo, M., 2018. Successful classification of experimental bone surface modifications (BSM) through machine learning algorithms: a solution to the controversial use of BSM in paleoanthropology? Archaeological and Anthropological Sciences. https://doi.org/10.1007/s12520-018-0684-9.
Domínguez-Rodrigo, M., Alcalá, L., 2016. 3.3-Million-Year-Old stone tools and butchery traces? More evidence needed. PaleoAnthropology. https://doi.org/10.4207/PA.2016.ART99.
Domínguez-Rodrigo, M., Baquedano, E., 2018. Distinguishing butchery cut marks from crocodile bite marks through machine learning methods. Sci. Rep. 8. https://doi.org/10.1038/s41598-018-24071-1.
Domínguez-Rodrigo, M., Barba, R., 2006. New estimates of tooth mark and percussion mark frequencies at the FLK zinj site: the carnivore-hominid-carnivore hypothesis falsified. J. Hum. Evol. 50 (2), 170–194.
Domínguez-Rodrigo, M., Barba, R., Egeland, C.P., 2007. Deconstructing Olduvai. Springer, The Netherlands.
Domínguez-Rodrigo, M., de Juana, S., Galán, A.B., Rodríguez, M., 2009. A new protocol to differentiate trampling marks from butchery marks. J. Archaeol. Sci. 36 (12), 2643–2654.
Domínguez-Rodrigo, M., Pickering, T.R., Bunn, H.T., 2011. Experimental study of cut marks made with rocks unmodified by human flaking and its bearing on claims of ~3.4-Million-Year-Old butchery evidence from dikika, Ethiopia. J. Archaeol. Sci. 39, 205–214. https://doi.org/10.1016/j.jas.2011.03.010.
Domínguez-Rodrigo, M., Saladié, P., Cáceres, I., Huguet, R., Yravedra, J., Rodríguez-Hidalgo, A., Martín, P., Pineda, A., Marín, J., Gené, C., Aramendi, J., Cobo-Sánchez, L., 2017. Use and abuse of cut mark analyses: the rorschach effect. J. Archaeol. Sci. 86, 14–23.
Domínguez-Rodrigo, M., Saladié, P., Cáceres, I., Huguet, R., Yravedra, J., Rodríguez-Hidalgo, A., Martín, P., Pineda, A., Marín, J., Gené, C., Aramendi, J., Cobo-Sánchez, L., 2018. Spilled Ink Blots the Mind: a Reply to Merrit et al. (2018) on Subjectivity and Bone Surface Modifications. J. Archaeol. Sci. https://doi.org/10.1016/j.jas.2018.09.003.
Dryden, I.L., Mardia, K.V., 1998. Statistical Shape Analysis. John Wiley & Sons, Chichester.
Duches, R., Nannini, N., Fontana, A., Boschin, F., Crezzini, J., Bernadini, F., Tuniz, C., Dalmeri, G., 2018. Archaeological bone injuries by lithic backed projectiles: new evidence on bear hunting from the late epigravettian site of cornafessa rock shelter (Italy). Archaeological and Anthropological Sciences. https://doi.org/10.1007/s12520-018-0674-y.
Egeland, C.P., Domínguez-Rodrigo, M., Pickering, T.R., Menter, C.G., Heaton, J.L., 2018. Hominin skeletal Part Abundances and claims of deliberate disposal of corpses in the middle pleistocene. Proc. Natl. Acad. Sci. Unit. States Am. https://doi.org/10.1073/pnas.1718678115.
Eulitz, M., Reiss, G., 2015. 3D reconstruction of SEM images by use of optical photogrammetry software. J. Struct. Biol. 191, 190–196. https://doi.org/10.1016/j.jsb.2015.06.010.
Fernández-Marchena, J.L., García-Argudo, G., Pedergnana, A., Valverde, I., 2016. Líneas, Manchas y Cía. Pautas Metodológicas para una Adecuada Interpretación Funcional. In: Conference: IX Jornadas de Jóvenes en Investigación Arqueológica, pp. 241–250 Santander 8-11 June, 2016.
Fiorillo, A.R., 1984. An experimental study of trampling: implications for the fossil record. In: Bonnichsen, R., Sorg, M.H. (Eds.), Bone Modification. University of Maine Press, Maine, pp. 73–98.
Fisher, J.W., 1995. Bone surface modification in zooarchaeology. J. Archaeol. Method Theor 2 (1), 7–68.
Fuentes-Sánchez, D., Galindo-Pellicena, M.Á., García-González, R., Carretero, J.M., Arsuaga, J.L., 2017. Experimental cut marks characterization using a confocal laser profilometer. In: Alonso, R., Baena, J., Canales, D. (Eds.), Playing with the Time. Experimental Archaeology and the Study of the Past, pp. 171–176.
Gifford-Gonzalez, D., 1991. Bones are not enough: analogues, knowledge, and interpretive strategies in zooarchaeology. J. Anthropol. Archaeol. 10, 215–254.
González-Aguilera, D., López Fernández, L., Rodríguez-González, P., Guerrero, D., Hernandez-Lopez, D., Remondino, F., Menna, F., Nocerino, E., Toschi, I., Ballabeni, A., Gaiani, M., 2016. Development of an all-purpose free photogrammetric tool. Conference: the International Archives of the Photogrammetry. Remote Senting and Spatial Information Sciences. https://doi.org/10.5194/isprsarchives-XLI-B6-31-2016.
Gümrükçu, M., Pante, M.C., 2018. Assessing the effects of fluvial abrasion on bone surface modifications using high-resolution 3-D scanning. J. Archaeol. Sci.: Report 21, 208–221.
Hervé, M., 2018. Package 'RVAideMemoire'. R package version 0.9. https://cran.r-project.org/package=RVAideMemoire.
Hutton, J., 1794. Experience Considered as the Cause of Knowledge, or as an Operation in Which the Human Intellect Is Made to Proceed, an Investigation of the Principles of Knowledge: and of the Progress of Reason, from Sense to Science and Philosophy 2. Strahan T Cadell, London, pp. 105–120 3.
Jordá Pardo, J., Cacho, C., 2013. Radiocarbono y Cronoestratigrafía del Registro Arqueológico Pleistoceno de La Peña de Estebanvela (Ayllón, Segovia, España). In: Cacho, C. (Ed.), Ocupaciones Magdalenienses en la Peña de Estebanvela. Junta de Castilla y León – CSIC, Madrid, pp. 75–92 (Coord.).
Lyell, C., 1830. Principles of Geology: Being an Attempt to Explain the Former Changes of the Earth's Surface, by Reference to Causes Now in Operation. Vertebrate Taphonomy. Cambridge University Press, Cambridge Penguin Classics, Wiltshire.
Malassé, A.D., Moigne, A.M., Singh, M., Calligaro, T., Karir, B., Gaillard, C., Kaur, A., Bhardwaj, V., Pal, S., Abdessadok, S., Sao, C.C., Gargani, J., Tudryn, A., Sanz, M.G., 2016. Intentional cut marks on bovid from the qurunwala zone, 2.6 ma, siwalik frontal range, northwestern India. Comptes Rendus Palevol 15, 317–339.
Marciani, G., Arrighi, S., Aureli, D., Spagnolo, V., Boscato, P., Ronchitelli, A., 2018. Middle palaeolithic lithic tools: techno-functional and use-wear analysis of target objects from SU 13 at the oscurusciuto rock shelter, southern Italy. Journal of Lithic Studies 5 (2). https://doi.org/10.2218/jls.2745.
Marín-Monfort, M.D., Pesquero, M.D., Fernández-Jalvo, Y., 2013. Compressive marks from gravel substrate on vertebrate remains: a preliminary experimental study. Quat. Int. https://doi.org/10.1016/j.quaint.2013.10.028.
Martín-Viveros, J.I., 2016. Microscopia digital, óptica y electrónica de barrido. Un enfoque complementario para el análisis de huellas de uso en industrias de sílex del Paleolítico medio. El nivel M del Abric Romaní (Capellades, Barcelona, España). Master's Thesis. Universitat Rovira i Virgili.
Martín-Viveros, J.I., Ollé, A., 2018. Use-wear and residue mapping in traceology. A multiscalar approach combining digital 3D, optical light and scanning Electron microscopy. In: Book of Abstracts XVIII° CONGRES UISPP Paris June 2018, pp. 875.
Martín-Viveros, J.I., Ollé, A., Chacón, M.G., 2018. Use-wear analysis of a specific mobile toolkit from the middle paleolithic site of abric romaní (barcelona, Spain). A complementary approach using optical light, digital 3D and scanning Electron microscopy. In: Book of Abstracts XVIII° CONGRES UISPP Paris June 2018. Paris, pp. 914–915.
Maté-González, M.A., Yravedra, J., González-Aguilera, D., Palomeque-González, J.F., Domínguez-Rodrigo, M., 2015. Microphotogrammetric characterization of cut marks on bones. J. Archaeol. Sci. 62, 128–142.
Maté-González, M.A., Palomeque-González, J.F., Yravedra, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2016. Micro-photogrammetric and morphometric differentiation of cut marks on bones using metal knives, quartzite and flint flakes. Journal of Archaeological and Anthropological Science. https://doi.org/10.1007/s12520-

016-0401-5.
Maté-González, M.A., Yravedra, J., Martín-Perea, D., Palomeque-González, J., San-Juan-Blazquez, M., Estaca-Gómez, V., Uribelarrea, D., Álvarez-Alonso, D., Cuartero, F., González-Aguilera, D., Domínguez-Rodrigo, M., 2017a. Flint and quartzite: distinguishing raw material through bone cut marks. Archaeometry. https://doi.org/10.1111/arcm.12327.
Maté-González, M.A., Aramendi, J., Yravedra, J., Blasco, R., Rosell, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2017b. Assessment of statistical agreement of three techniques for the study of cut marks: 3D digital microscope, laser scanning confocal microscopy and micro-photogrammetry. J. Microsc. 267 (3), 356–370. https://doi.org/10.1111/jmi.12575.
Maté-González, M.A., Aramendi, J., Yravedra, J., González-Aguilera, D., 2017c. Statistical comparison between low-cost methods for 3D characterization of cut-marks on bones. Rem. Sens. https://doi.org/10.3390/rs9090873.
McPherron, S.P., Alemseged, Z., Marean, C.W., Wynn, J.G., Reed, D., Geraads, D., Bobe, R., Béarat, H.A., 2010. Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at dikika, Ethiopia. Nature 466, 857–860.
Merton, R.K., 1967. On Theoretical Sociology: Five Essays, Old and New. The Free Press, New York.
Moclán, A., Huguet, R., Márquez, B., Domínguez-Rodrigo, M., Gómez-Miguelsanz, C., Vergès, J.M., Laplana, C., Arsuaga, J.L., Pérez-González, A., Baquedano, E., 2018. Cut Marks Made with Quartz Tools: an Experimental Framework for Understanding Cut Mark Morphology, and its Use at the Middle Palaeolithic Site of the Navalmaíllo Rock Shelter (Pinilla del Valle, Madrid, Spain). Quat. Int. https://doi.org/10.1016/j.quaint.2018.09.033.
Moretti, E., Arrighi, S., Boschin, F., Crezzini, J., Aureli, D., Ronchitelli, A., 2015. Using 3D microscopy to analyze experimental cut marks on animal bones produced with different stone tools. Ethnobiology Letters 6 (2), 14–22. https://doi.org/10.14237/ebl.6.1.2015.349.
Öhman, C., Zwierzak, I., Beleani, M., Viceconti, M., 2012. Human bone hardness seems to depend on tissue type but not on anatomical site in the Long bones of an old subject. Journal of Engineering in Medicine 227 (2), 200–206.
Oliver, J.S., 1984. Analogues and site context: bone damages from shield trap cave (24CB91), carbon county, Montana, USA. In: Bonnischen, R., Sorg, M.H. (Eds.), Bone Modification. University of Maine Press, Maine, pp. 61–72.
Olsen, S.L., 1988. The identification of stone and metal tool marks on bone artefacts. Scanning Electron microscopy in archaeology. BAR Int. Ser. 452, 337–360.
Olsen, S.L., Shipman, P., 1988. Surface modification on bone: trampling versus butchery. J. Archaeol. Sci. 15, 535–553.
Orlikoff, E.R., Keevil, T.L., Pante, M.C., 2018. Quantitiative analysis of the micro-morphology of trampling-induced abrasion and stone tool cut marks on bone surfaces. In: Conference: Paleoanthropology Society Conference, https://doi.org/10.13140/RG.2.2.31360.94721. 2018.
Otárolla-Castillo, E., Torquato, M., Hawkins, H.C., James, E., Harries, J.A., Marean, C.W., McPherron, S.P., Thompson, J.C., 2017. Differentiating between cutting actions on bone using 3D geometric morphometrics and bayesian analyses with implications to human evolution. J. Archaeol. Sci. https://doi.org/10.1016/j.jas.2017.10.004.
Oxilia, G., Peresani, M., Romandini, M., Matteucci, C., Spiteri, C.D., Henry, A., Shulz, D., Archer, W., Crezzini, J., Boschin, F., Boscato, P., Jaouen, K., Dogandzic, T., Brolio, A., Moggi-Cecchi, J., Fiorenza, L., Hublin, J.J., Kullmer, O., Benazzi, S., 2015. Earliest evidence of dental caries manipulation in the late upper palaeolithic. Sci. Rep. 5. https://doi.org/10.1038/srep12150.
Pante, M.C., Muttart, M.V., Keevil, T.L., Blumenschine, R.J., Njau, J.K., Merritt, S.R., 2017. A new high-resolution 3-D quantitative method for identifying bone surface modifications with implications for the Early Stone Age archaeological record. J. Hum. Evol. 102, 1–11.
Pérez, L., Sanchis, A., Hernández, C.M., Galván, B., Sala, R., Mallol, C., 2017. Hearths and bones: an experimental study to explore temporality in archaeological contexts based on taphonomical changes in burnt bones. J. Archaeol. Sci.: Report 11, 287–309.
Pineda, A., Cáceres, I., Saladié, P., Huguet, R., Rosas, A., Vallverdú, J., Under revision. Tumbling effects and its application to archaeological deposits: the case of barranc de la Boella (tarragona, Spain). Submitted to Journal of Archaeological Science, Under Review.
Pineda, A., Saladié, P., Vergés, J.M., Huguet, R., Cáceres, I., Vallverdú, J., 2014. Trampling *versus* cut marks on chemically altered surfaces: an experimental approach and archaeological application at the barranc de la Boella site (la Canonja, tarraona, Spain). J. Archaeol. Sci. 50, 84–93.
Playfair, J., 1802. Figure of the Earth, Illustrations of the Huttonian Theory of the Earth. William Creech, Edinburgh, pp. 488–509.
Potts, R., Shipman, P., 1981. Cutmarks made by stone tools on bones from olduvai gorge, Tanzania. Nature 291, 577–1981.
Revedin, A., Aranguren, B., Becattinia, R., Longo, L., Marconi, E., Mariotti Lippi, M., Skakun, N., Sinitsyn, A., Spiridonova, E., Svoboda, J., 2010. Thirty thousand-year-old evidence of plant food processing. Proc. Natl. Acad. Sci. Unit. States Am. 107 (44), 18815–18819.
Reynard, J.P., 2014. Trampling in coastal sites: an experimental study on the effects of shell on bone in coastal sediment. Quat. Int. 330, 156–170.
Rodríguez-Hidalgo, A., Morales, J.I., Cebriá, A., Courtenay, L.A., Fernández-Marchena, J.L., García-Argudo, G.G., Marín, J., Saladié, P., Soto, M., Tejero, J., Fullola, J., 2018. The châtelperronian Neandertals of Cova Foradada (Calafell, Spain) used Iberian imperial eagle phalanges for symbolic purposes. PeerJ Preprints 6, e27133v1. https://doi.org/10.7287/peerj.preprints.27133v1.
Rohlf, F.J., 1999. Shape statistics: Procrustes superimpositions and tangent spaces. J. Classif. 16 (2), 197–223.
Ronchitelli, A., Mugnaini, S., Arrighi, S., Atrei, A., Capecchi, G., Giamello, M., Longo, L., Marchettini, N., Viti, C., Moroni, A., 2015. When technology joins symbolic behaviour: the gravettian burials at grotta paglicci (rignano garganico – foggia – southern Italy). Quat. Int. 423–441 359-360.
Rufa, A., Blasco, R., Rosell, J., Vaquero, M., 2017. What is going on at the Molí del Salt site? A Zooarchaeological Approach to the Last Hunter-Gatherers from South Catalonia. Hist. Biol. https://doi.org/10.1080/08912963.2017.1315685.
Sahle, Y., El Zaatari, S., White, T.D., 2017. Hominid butchers and biting crocodiles in the african plio-pleistocene. Proc. Natl. Acad. Sci. Unit. States Am. https://doi.org/10.1073/pnas.1716317114.
Schiffer, M.B., 1987. Formation Processes of the Archaeological Record. University of New Mexico Press.
Shipman, P., Rose, J., 1983. Evidence of butchery and hominid activities at torralba and ambrona; an evaluation using microscopic techniques. J. Archaeol. Sci. 10, 465–474.
Shipman, P., Fisher, D.C., Rose, J.J., 1984a. Mastodon butchery: microscopic evidence of carcass processing and bone tool use. Paleobiology 10 (3), 358–365.
Shipman, P., Foster, G., Schoeninger, M., 1984b. Burnt bones and teeth: an experimental study of color, morphology, crystal structure and shrinkage. J. Archaeol. Sci. 11, 307–325.
Slice, D.E., 2001. Landmark coordinates aligned by Procrustes analysis do not lie in kendall's shape space. Syst. Biol. 50 (1), 141–149. https://doi.org/10.1080/10635150119110.
Stemp, W.J., Watson, A.S., Evans, A.A., 2015. Surface Analysis of Stone and Bone Tools. Surface Topography: Metrology and Properties. https://doi.org/10.1088/2051-672X/4/1/013001.
Stinnesbeck, S.R., Stinnesbeck, W., Terrzas Mata, A., Avilés Olguín, J., Benavente Sanvicente, M., Zell, P., Frey, E., Lindauer, S., Sandoval, C.R., Velázquez Morlet, A., Acevez Nuñez, E., González González, A., 2018. The muknal cave near talum, Mexico: an early-holocene funeral site on the yucután peninsula. Holocene. https://doi.org/10.1177/0959683618798124.
Tafti, A.P., Kirkpatrick, A.B., Alavi, Z., Owen, H.A., Yu, Z., 2015. Recent advances in 3D SEM surface reconstruction. Micron 78, 54–66. https://doi.org/10.1016/j.micron.2015.07.005.
The Oxford English Dictionary, 2018. The Oxford English Dictionary. Oxford University Press, Oxford.
Thompson, J.C., McPherron, S., Bobe, R., Reed, D., Barr, A., Wynn, J., Marean, C.W., Geraads, D., Alemseged, Z., 2015. Taphonomy of fossils from the hominin-bearing deposits at dikika, ehtiopia. J. Hum. Evol. 86, 122–135.
Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer, New York. http://stats.ox.ac.uk/pub/MASS4.
Vergès, J.M., Morales, J.I., 2014. The gigapixel image concept for graphic SEM documentation. Applications in archaeological use-wear studies. Micron 65, 15–19. https://doi.org/10.1016/j.micron.2014.04.009.
Walden, S.J., Evans, S.L., Mulville, J., 2017. Changes in vickers hardness during the decomposition of bone: possibilities for forensic anthropology. Journal of Mechanical Behavior of Biomedical Materials 65, 672–678.
Walker, P.L., Long, J.C., 1977. An experimental study of morphological characteristics of tool marks. Am. Antiq. 42 (4), 605–616.
Whewell, W., 1847. Of the Fundamental Antithesis of Philosophy, the Philosophy of the Inductive Sciences Founded upon Their History 2. John W Parker, London, pp. 16–50.
Wierer, U., Arrighi, S., Bertola, S., Kaufmann, G., Baumgarten, B., Pedrotti, A., Pernter, P., Pelegrin, J., 2018. The iceman's lithic toolkit: raw material, technology, typology and use. PLoS One 13 (6). https://doi.org/10.1371/journal.pone.0198292.
Yravedra, J., 2005. Patrones de Aprovechamiento de Recursos Animales en el Pleistoceno Superior de la Península Ibérica: Estudio Tafonómico y Zooarqueológico de los Yacimientos del Esquilleu, Amalda, Cueva Ambrosio y la Peña de Estebanvela. Doctoral Thesis. Universidad Nacional de Educación a Distancia (UNED), Madrid.
Yravedra, J., García Vargas, E., Maté González, M.A., Aramendi, J., Palomeque-González, J., Vallés-Iriso, J., Matasanz-Vicente, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2017a. The use of Micro-Photogrammetry and Geometric Morphometrics for identifying carnivore agency in bone assemblage. Journal of Archaeological Science Reports 14, 106–115.
Yravedra, J., Maté-González, M.A., Palomeque-González, J.F., Aramendi, J., Estaca-Gómez, V., Blazquez, M.S., García Vargas, E., Organista, E., González-Aguilera, D., Arriaza, M.C., Cobo-Sánchez, L., Gidna, A., Uribelarrea del Val, D., Baquedano, E., Mabulla, A., Domínguez-Rodrigo, M., 2017b. A new approach to raw material use in the exploitation of animal carcasses at BK (upper bed II, olduvai gorge, Tanzania): a micro-photogrammetric and geometric morphometric analysis of fossil cut marks. Boreas. https://doi.org/10.1111/bor.12224.
Yravedra, J., Diez-Martín, F., Egeland, C.P., Maté-González, M.Á., Palomeque-González, J.F., Arriaza, M.C., Aramendi, J., García Vargas, E., Estaca-Gómez, V., Sánchez, P., Fraile, C., Duque, J., de Francisco Rodríguez, S., González-Aguilera, D., Uribelarrea, D., Mabulla, A., Baquedano, E., Domínguez-Rodrigo, M., 2017c. FLK-West (Lower Bed II, Olduvai Gorge, Tanzania): a newearly Acheulean site with evidence for human exploitation of fauna. Boreas. https://doi.org/10.1111/bor.12243. ISSN 0300-9483.
Yravedra, J., Andrés-Chaín, M., Martos, J.A., Marquer, L., Avezuela, B., Jordá-Pardo, J., Martín-Lema, Sesé, C., Valdivia, J., 2018a. Recurrent magdalenian occupation in the interior of the iberian peninsula: new insights from the archaeological site of La Peña de Estebanvela (segovia, Spain). Archaeological and Anthropological Sciences. https://doi.org/10.1007/s12520-018-1620-z.
Yravedra, J., Aramendi, J., Maté-González, M.Á., Courtenay, L.A., González-Aguilera, D., 2018b. Differentiating percussion pits and carnivore tooth pits using 3D reconstructions and geometric morphometrics. PLoS One, e0194324. https://doi.org/10.1371/journal.pone.0194324.

Appendix 5
Publication Number 2

The second published article produced with data from this Master's Thesis.

Courtenay, L.A.; Yravedra, J.; Huguet, R.; Aramendi, J.; Maté-González, M.Á.; González-Aguilera, D.; Arriaza, M.C. (2019) Combining Machine Learning Algorithms and Geometric Morphometrics: a Study of Carnivore Tooth Marks, *Palaeogeography, Palaeoclimatology, Palaeoecology*. 522:28-29 DOI: 10.1016/j.palaeo.2019.03.007

Palaeogeography, Palaeoclimatology, Palaeoecology 522 (2019) 28–39

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

journal homepage: www.elsevier.com/locate/palaeo

Combining machine learning algorithms and geometric morphometrics: A study of carnivore tooth marks

Lloyd A. Courtenay[a,b,c,*], José Yravedra[c,d], Rosa Huguet[b,a,e], Julia Aramendi[c,f], Miguel Ángel Maté-González[g], Diego González-Aguilera[g], Mari Carmen Arriaza[h,i]

[a] Àrea de Prehistòria, Universitat Rovira i Virgili (URV), Avinguda de Catalunya 35, 43002 Tarragona, Spain
[b] Institut de Paleoecologia Humana i Evolució Social (IPHES), Zona Educacional, Campus Sescelades URV (Edifici W3) E3, 43700 Tarragona, Spain
[c] Department of Prehistory, Complutense University, Prof. Aranguren s/n, 28040 Madrid, Spain
[d] C. A. I. Archaeometry and Archaeological Analysis, Complutense University, Professor Aranguren s/n, 28040 Madrid, Spain
[e] Unit Associated to CSIC, Departamento de Paleobiologia, Museo de Ciencias Naturales, calle José Gutiérrez Abascal, s/n, 28006 Madrid, Spain
[f] IDEA (Institute of Evolution in Africa), Covarrubias 36, 28010 Madrid, Spain
[g] Department of Cartography and Terrain Engineering, Polytechnic School of Avila, University of Salamanca, Hornos Caleros 50, 05003 Avila, Spain
[h] School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Private Bag 3, 2050, South Africa
[i] Centre of Excellence in Paleosciences, University of the Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Keywords:
Artificial Intelligence
Bone Surface modifications
Statistics
Taphonomy

ABSTRACT

Since the 1980s an intense scientific debate has revolved around the hunting capacities of early hominin populations and the behavioral patterns of carnivores sharing the same ecosystem, and thus competing for the same resources. This debate, commonly known as the hunter-scavenger debate, fostered the emergence of a new research line into the Bone Surface Modifications (BSMs) produced by both taphonomic agents. Throughout the following 20 years, multiple studies concerning the action of carnivores have been developed, with a particular focus on the oldest archaeological sites in East Africa. Recent technological advances applied to taphonomy have provided new insight into carnivore BSMs. A newly developed part of this work relies on Geometric Morphometrics (GMM) studies aimed at discerning carnivore agency through the morphologic characterization of tooth scores and pits. GMM studies have produced promising results, however methodological limitations are still present. This paper presents the first combined application of Machine Learning (ML) algorithms and GMM to the analysis of carnivore tooth marks, generating classification rates of 100% between carnivore species in some cases.

1. Introduction

A fundamental line of research in human evolution consists in understanding early hominin subsistence patterns and the interactions with other paleo-communities in the paleoecosystem. Studying such a complex ecological niche ca. 3 to 1 Ma has fueled a wide range of debates regarding early human behavior, mostly focused on both East and South African sites. The famous *osteodontokeratic* culture was associated with *Australopiths* remains from the Makapansgat (South Africa) Member 3 deposit in the mid-20th century (Dart, 1957). Subsequently, Brain's (1967, 1969) neotaphonomic research showed that bone assemblages produced after human butchery and carnivore ravaging was very similar to that found in the Makapansgat assemblage. More importantly, the presence of carnivore activity was identified at the Cradle of Humankind sites in South Africa (Kromdraai, Swartkrans and Sterkfontein) on hominin remains suggesting that *Australopiths* carcasses were accumulated by carnivores (Brain, 1981). From this perspective *Australopiths* were consequently depicted as prey rather than predators (Brain, 1981).

Similarly, questions regarding the subsistence patterns of early hominins became a heated source of discussion towards the end of the 20th century based on the study of East African archaeological sites such as Koobi Fora or Olduvai Gorge (Binford, 1981; Bunn, 1981). Throughout the 1980s and 1990s, the *hunter-scavenger* debate confronted the distinction of different taphonomic agents involved in bone modification, so that discerning between anthropogenic and carnivore activity became a key practice in archaeological research. Part of this development focused greatly on the ecology of different carnivore species and ethnoarchaeological studies, in which taphonomy played a pivotal role (Bunn, 1981; Blumenschine, 1988; Blumenschine and

* Corresponding author at: Àrea de Prehistòria, Universitat Rovira i Virgili (URV), Avinguda de Catalunya 35, 43002 Tarragona, Spain.
E-mail address: ladc1995@gmail.com (L.A. Courtenay).

https://doi.org/10.1016/j.palaeo.2019.03.007
Received 5 January 2019; Received in revised form 5 March 2019; Accepted 5 March 2019
Available online 07 March 2019
0031-0182/

Marean, 1993; Selvaggio, 1994; Blumenschine, 1995; Capaldo, 1997; Stanford and Bunn, 2001; Domínguez-Rodrigo, 2002; Pickering and Bunn, 2007; Domínguez-Rodrigo et al., 2007), expanding on previous studies addressing this matter (Dawkins, 1863; Schaller and Lowther, 1969; Sutcliffe, 1970; Houtson, 1979).

An important amount of taphonomic studies have tried to characterize the different types of osteological accumulations produced by carnivores, mainly based on the Bone Surface Modifications (BSMs) produced. Through this, the action of hyenas (Sutcliffe, 1970; Brain, 1981; Cruz-Uribe, 1991; Domínguez-Rodrigo, 2001; Pickering, 2002; Egeland et al., 2008; Domínguez-Rodrigo and Pickering, 2010; Domínguez-Rodrigo et al., 2015), lions (Domínguez-Rodrigo, 1999; Domínguez-Rodrigo et al., 2012; Gidna et al., 2013, 2014; Arriaza et al., 2016), wolves (Haynes, 1983a, b; Willey and Snyder, 1989; Stiner, 2004; Yravedra et al., 2011, 2012) and leopards (Brain, 1981; Kerbis-Peterhans, 1990; Ruiter and Berger, 2000; Pickering et al., 2011) have been widely studied. Additionally, studies of jaguars have also received attention, yet to a lesser degree (Martín, 2008; Domínguez-Rodrigo et al., 2015).

Multiple criteria have been analyzed to determine the agents involved in bone modification at archaeological sites. These include the type of prey (Brain, 1981; Haynes, 1982), skeletal part profiles (Marean and Frey, 1997; Bartram and Marean, 1999; Pickering et al., 2003; Faith et al., 2007), prey mortality profiles (Stiner, 1994; Bunn and Pickering, 2010), and tooth marks (Selvaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003; Delaney-Rivera et al., 2009; Andrés et al., 2012; Domínguez-Rodrigo et al., 2012). Tooth marks can further be divided into different categories, produced through different masticatory actions during carcass consumption. Among these, the two most frequent types are pits and scores. While pits present a circular morphology created by the direct imprint of the carnivore's tooth, a score is a shallow elongated groove across the surface of the bone with an equally rounded base (Binford, 1981; Blumenschine, 1995).

Initial studies into these two BSMs focused primarily on biometric dimensional properties (Selvaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003). While using 95% confidence intervals have been able to reduce a great percentage of equifinality present through studies of this type (Andrés et al., 2012), overlapping of samples is still large and, in many cases, little information serves to accurately build hypotheses. On the other hand, other authors have studied the properties of bone fracture patterns, differentiating carnivore agency from anthropogenic activities via fracture plane angles (Capaldo and Blumenschine, 1994; Pickering et al., 2005; Pickering and Egeland, 2006; Alcántara et al., 2006; Galán et al., 2009; Moclán and Domínguez-Rodrigo, 2018). Development of new variables such as the taphotypes have successfully differentiated between spotted hyenas and lions, alongside different types of felids (Domínguez-Rodrigo et al., 2015), nevertheless, the identification of different carnivore taxa involved in bone modification remains controversial.

The application of Geometric Morphometrics (GMM) to taphonomy has transformed the analysis of both anthropogenic (Maté-González et al., 2015, 2016, 2017a; Courtenay et al., 2017), and carnivore BSMs, including tooth scores (Yravedra et al., 2017), and tooth pits (Aramendi et al., 2017a, Arriaza et al., 2018). Despite the apparent success of these studies, certain aspects still need to be improved. Aramendi et al. (2017a), Arriaza et al. (2017, 2018) and Yravedra et al. (2017) were able to determine the action of some carnivore species based on the morphology of their traces up to certain degree, not reaching 100% of certainty. Additionally, while the tooth marks of hyenas and lions were distinguishable, wolves and jaguars are much harder to define. The limitations imposed by sample size and the number of carnivore species included in the analysis might have also affected the resolution obtained with these methods.

The pioneering introduction of Machine Learning (ML) algorithms into archaeological research, however, offers an excellent opportunity to analyse BSMs with a higher precision. The first introduction of ML methods to the study of carnivore activity, based on the skeletal part representation observed in modern carnivore bone accumulations, provided 100% classification rates in some cases (Arriaza and Domínguez-Rodrigo, 2016). Equally remarkable results have been presented for the analysis of anthropogenic BSMs by means of ML techniques (Domínguez-Rodrigo, 2018; Domínguez-Rodrigo and Baquedano, 2018), as well as fracture patterns (Moclán and Domínguez-Rodrigo, Under Revision).

2. Materials and methods

In this paper we present the first combined effort to join GMM analysis and ML techniques to discern carnivore agency based on tooth mark morphology, inspired by implementations of advanced statistical techniques in taphonomy. For the purpose of this study we have reassessed the GMM models proposed by Aramendi et al. (2017a) and Yravedra et al. (2017), implementing ML algorithms for the processing of morphometric data.

2.1. Tooth mark samples

A total of 89 carnivore tooth pits and 127 scores on adult horse long bones were compared. All bones contained flesh when presented to the individuals. These included tooth marks generated by spotted hyenas (pits = 21, scores = 33), jaguars (pits = 20, scores = 34) and lions (pits = 24, scores = 30) in a controlled setting, in the Cabárceno Nature Park in Cantabria (Spain), were studied. All the marks were produced by multiple (minimum 2) adult individuals for each species. Tooth marks produced by wolves (pits = 24) were obtained from Cabárceno Park, as well as the natural wolf sites in mount Campelo, near Sobrado Dos Montxes, Galicia (scores = 30). In the case of wolves, the number of wolves ranged between 10 and 15 wild individuals.

Only tooth marks on long bone shafts (namely tibiae and radii) were selected for two main reasons. Firstly, diaphyses are denser than epiphyses, therefore teeth tend to penetrate cortical layers less. Secondly, diaphyses present a higher survival rate, thus the use of a sample based on shafts would be a more useful framework for future archaeo-paleontological analogies. Bone epiphyses are usually more susceptible to being destroyed during carnivorous feeding (e.g. furrowing) or taphonomically affected through fluvial agents (Lyman, 1994).

Pits and scores were selected on the basis of their preservation and general conditions. Inconspicuous or superficial tooth marks and those that present a bad cortical preservation or some type of post-depositional alteration were excluded from the analysis.

For more details about the lion sample see Gidna et al. (2013), for the hyena and jaguar samples see Domínguez-Rodrigo et al. (2015), and for wolf pit and score samples see Yravedra et al. (2011, 2012, 2017).

2.2. Virtual reconstruction of marks

Micro-Photogrammetry and computer vision techniques were used to create high-resolution 3D models of the pits and scores (Fig. 1). Precise metrical models were generated using images taken with oblique photography using a CANON EOS 700D with a 60-mm macro lens (Table S1) and following the specified protocol explained in Maté-González et al. (2015). The camera was self-calibrated to simultaneously compute the interior and exterior camera parameters. For data collection, a total of 9–10 photos were taken for each mark. The number of photos varies depending on the geometry of the bone and the shape of the mark (Aramendi et al., 2017a; Yravedra et al., 2017). Once the photographs had been taken, they were processed so as to generate a 3D model for each mark with the open-source photogrammetric reconstruction software GRAPHOS (inteGRAted PHOtogrammetric Suite) (González-Aguilera et al., 2016).

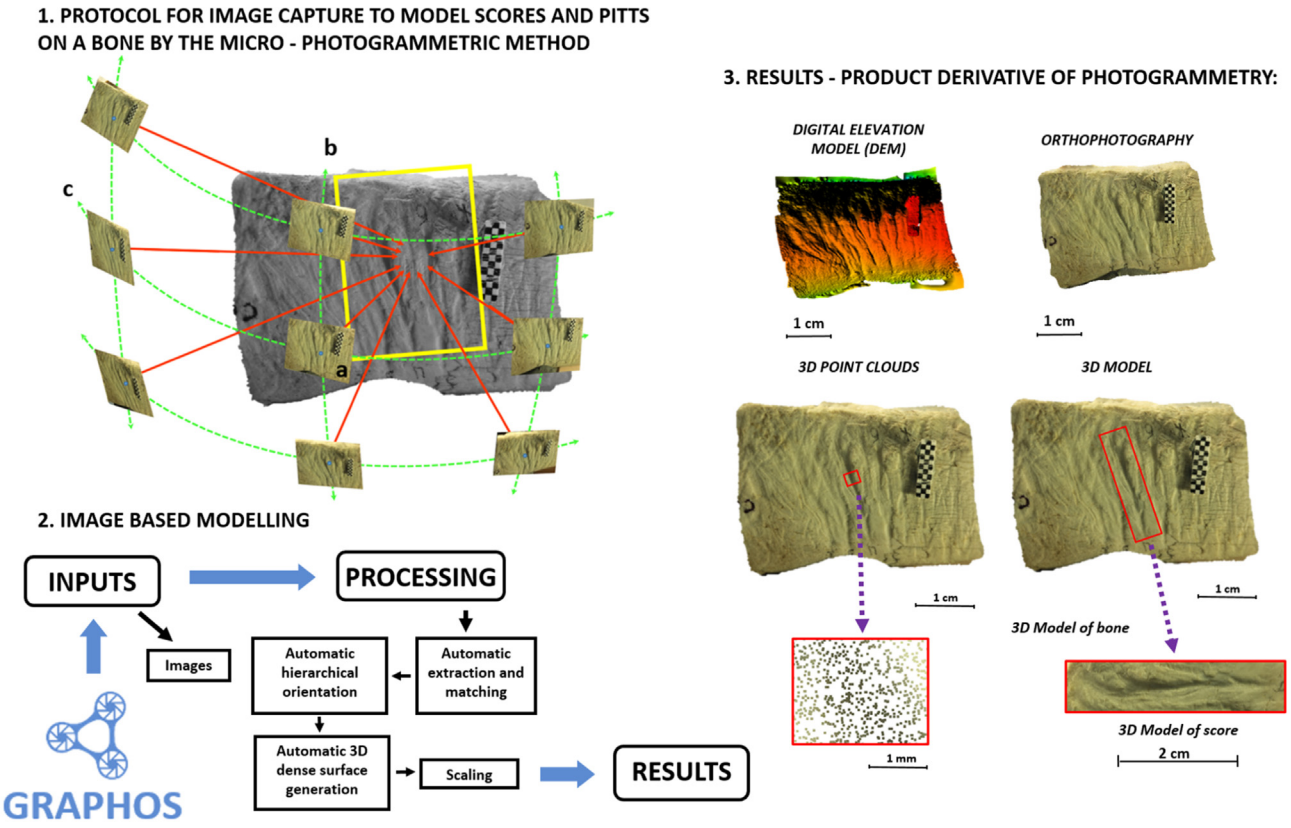


Fig. 1. Protocol for Image Capture and 3D reconstruction for each tooth mark prior to processing (Maté-González et al., 2015; Yravedra et al., 2017).

2.3. Geometric morphometrics

Geometric morphometric analyses of the tooth marks were based on two different models. A 7 2D homologous landmarks, as described by Yravedra et al. (2017), was used on tooth scores while tooth pits were studied using 17 3D homologous landmarks, as described by Aramendi et al. (2017a). Both models can be consulted in Table 1. All statistical analyses were carried out in the free software R (www.rproject.org, Core-Team, 2015). All R packages and libraries with their respected bibliography can be consulted in Supplemental Appendix 1.

Tooth score cross sections were obtained through importing the 3D digital reconstruction of each tooth mark into a Global Mapper software (Fig. 2), where marks could be treated, extracting sections at mid length (between 30% and 70% of the marks' length) (Maté-González et al.,

Table 1
Geometric morphometric landmark models used for comparisons between taxa. Including Yravedra et al.'s (2017) 2D 7-lanmdark model and Aramendi et al. (2017a) 3D 17-landmark model.

Model	N°	Landmark	Description
2D	1	Left shoulder	Upper limit of the left side of the mark
	2	Midpoint left wall	Halfway between the bottom of the mark and the left shoulder
	3	Base left wall	Point where the left wall and bottom of the mark meet
	4	Depth	Deepest point of the mark
	5	Base right wall	Point where the right wall and bottom of the mark meet
	6	Midpoint right wall	Halfway between the bottom of the mark and the right shoulder
	7	Right shoulder	Upper limit of the right side of the mark
3D	1	Length A	Upper limit of the longitudinal axis
	2	Length B	Lower limit of the longitudinal axis
	3	Width A	Left limit of the breadth axis
	4	Width B	Right limit of the breath axis
	5	Depth	Most centered lowest point of the pit
	6	Left upper half A	Point at the first third between the upper limit of the long axis and the left limit of the breadth axis
	7	Left upper half B	Point at the second third between the upper limit of the long axis and the left limit of the breadth axis
	8	Left lower half A	Point at the first third between the left limit of the breadth axis and the lower limit of the long axis
	9	Left lower half B	Point at the second third between the left limit of the breadth axis and the lower limit of the long axis
	10	Right upper half A	Point at the first third between the upper limit of the long axis and the right limit of the breadth axis
	11	Right upper half B	Point at the second third between the upper limit of the long axis and the right limit of the breadth axis
	12	Right lower half A	Point at the first third between the right limit of the breadth axis and the lower limit of the long axis
	13	Right lower half B	Point at the second third between the right limit of the breadth axis and the lower limit of the long axis
	14	Interior length A	Upper inflection point on the longitudinal axis
	15	Interior length B	Lower inflection point on the longitudinal axis
	16	Interior width A	Left inflection point on the breadth axis
	17	Interior width B	Right inflection point on the breadth axis

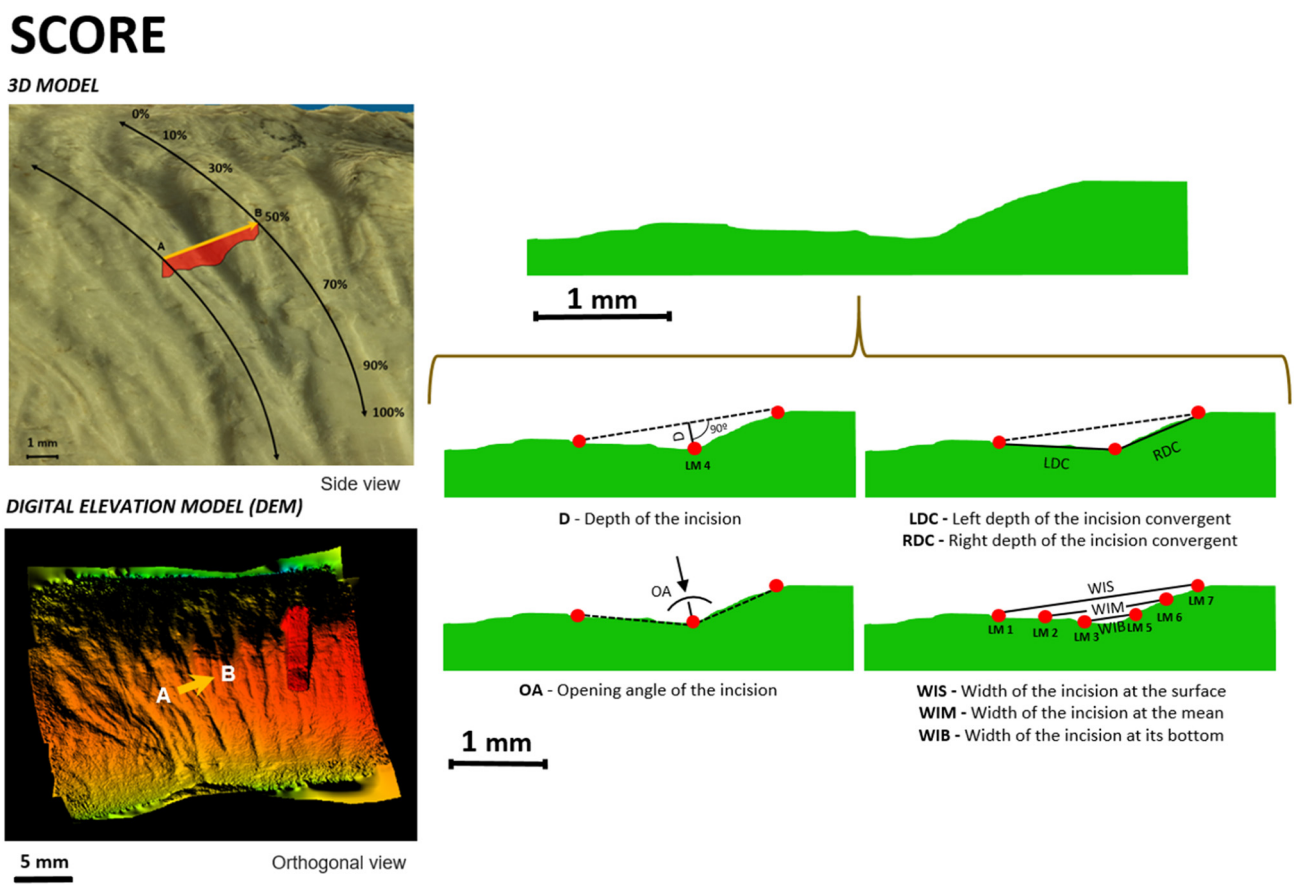


Fig. 2. Location of the seven landmarks in the 2D Geometric Morphometric model for tooth score cross-sections (Yravedra et al., 2017).

2015; Arriaza et al., 2017; Yravedra et al., 2017). These profiles were then extracted and imported into the free software tpsDig2 (v.2.1.7) where landmark data was collected.

3D digital reconstructions of each tooth pit were exported into the Avizo software (Visualization Sciences Group, USA) for landmarking (Fig. 3). Files containing landmark coordinates were then edited and imported into R where a full Procrustes fit and an orthogonal tangent projection (Dryden and Mardia, 1998) were used to normalise the data for further multivariate statistical analyses. This technique, commonly referred to as Generalized Procrustes Analysis (GPA), is used to standardize the form information through the application of superimposition procedures including translation, rotation and scaling. Any remaining differences are exposed through patterns of variation and covariation that can be assessed through several statistical tests (Slice, 2001; Rohlf, 1999). Principal component analyses (PCA) in shape and form space on the Procrustes superimposed landmarks were performed to reduce the large sets of variables to fewer dimensions that represent most of overall variance (Bookstein, 1991; Mitteroecker and Gunz, 2009). The PC scores calculated by the PCAs were then extracted and used for ML training.

2.4. Machine learning algorithms

Recent applications of ML algorithms in archaeology have been described as powerful classifiers (Domínguez-Rodrigo, 2018). ML provides a treatment of raw data using structured techniques that assign meaning to the data (Lantz, 2013). Much like in humans, the learning process consists in data retrieval (input), translation (abstraction) and a final output where the abstract data is used to form the basis behind the action (generalization) (Lantz, 2013). In this case, ML is particularly useful in handling large amounts of data and efficiently providing a means of handing this information; through classification.

To ensure the accuracy of training models, the samples were bootstrapped 1000 times. To train each algorithm and generate our classification models, 70% of the sample was separated. Validation was then carried out on the remaining 30% of the sample. This is a standard process in testing predictive model efficiency. Model evaluation is performed through resampling techniques that can be used to estimate performance. This can be carried out using multiple methods; in this case we preferred the method based on subsampling and submodeling through k fold cross validation ($k = 10$). One main advantage of ML algorithms is the possibility of tuning the parameters in order to improve performance. This can either be done manually via altering values such as *gamma* and *cost*, in the case of SVM models, or they can be tuned using additional lines of R coding (Supplemental Appendix). The parameters achieved through tuning are then applied in the models, thus improving performance.

A random selection of 9 common ML algorithms was selected. Much like in the case of GMM statistical analysis of variance, PC scores produced through PCA were used for ML analysis. Comparison of ML algorithm performance was then carried out considering Kappa, sensitivity, specificity and balanced accuracy values through confusion matrix tables. The Kappa statistic adjusts accuracy through considering the possibility of a correct prediction by change alone (Lantz, 2013). These values are presented as a value between −1 and 1, with any value above 0.8 considered as a powerful predictive model. Sensitivity and Specificity tests combine different evaluations of Type I and Type II statistical errors in proportion with the rest of the calculated confusion matrix. Calculation of these ratios is a common practice in medical statistics, whereby sensitivity defines the likelihood of correctly diagnosing an illness while specificity is defined by the likelihood of correctly identifying the healthiness of an individual (Fawcett, 2006).

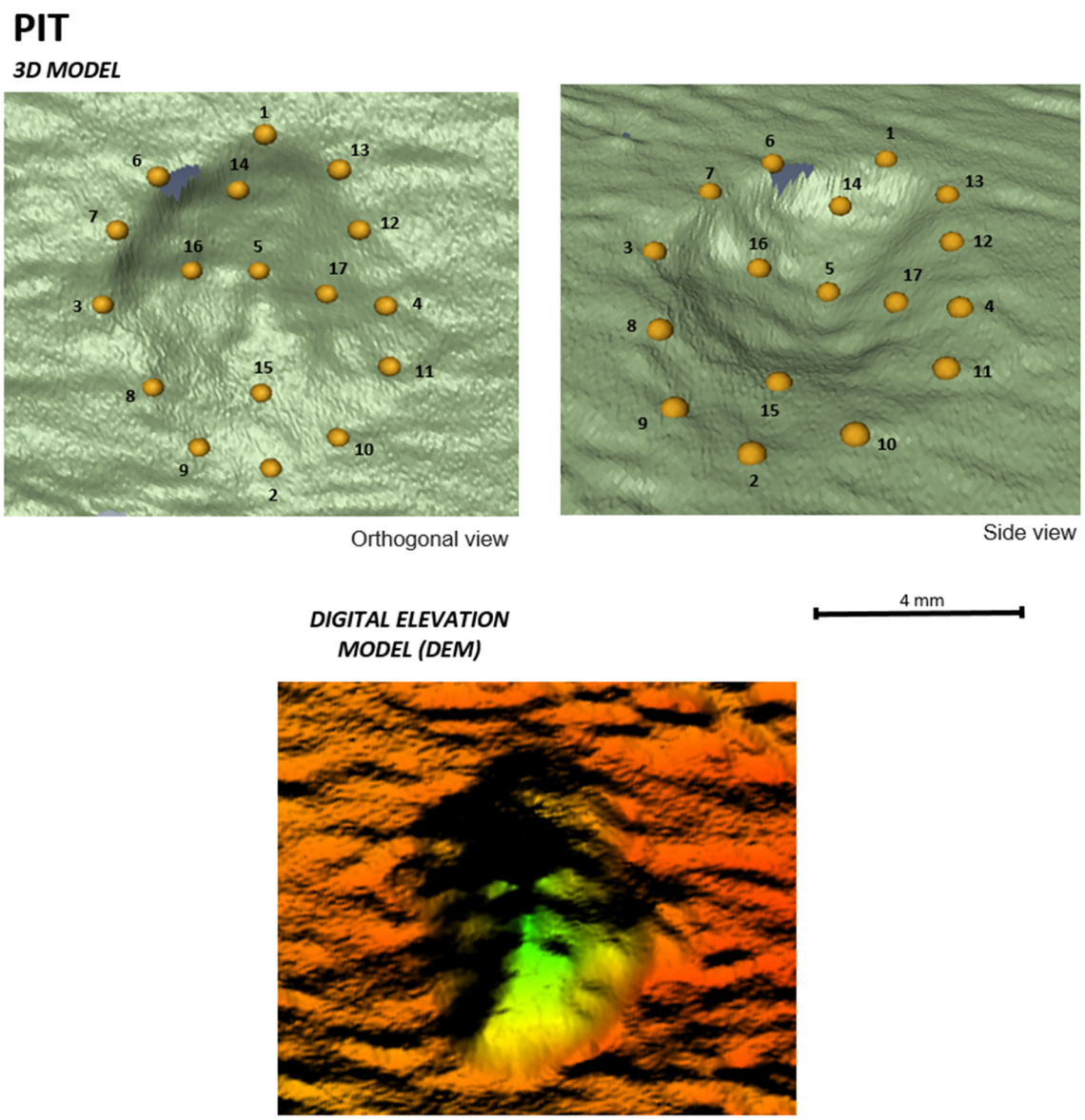


Fig. 3. Location of the seventeen landmarks in the 3D Geometric Morphometric model for tooth pits Aramendi et al. (2017a).

These can also be referred to as the true positive and true negative rates, thus defining the likeliness of misclassification (low number = high misclassification). A combination of these results can provide an additional balanced accuracy value that defines the power of the model (Lantz, 2013).

The 9 ML algorithms selected are the following:

Neural Network (NNET). NNET is inspired by human brain activities, using a computational model that mimics brain patterns. This method uses an interconnected group of nodes, involving input, connected weights, processing elements and a resulting output (Lantz, 2013; Wei and Chiu, 2015). Nodes are built through multiple regression methods. Training of the neural networks consists in the adjustment of weights through successive layers of nodes (Günther and Fritsch, 2010). The input data fed to the neural layers (perceptrons) is transformed via specific nonlinear sigmoidal functions. The parameters of these functions are usually optimised to minimize SQR (Sum of Square Residuals). These parameters exhibit a tendency to overfit the training data set. To avoid this, weight decay is used to reduce the model error for a given value of lambda (Domínguez-Rodrigo, 2018). This λ parameter must be specified together with the number of hidden units.

Support Vector Machines (SVM). SVM maps input data into high dimensional feature space defined by the *kernel* function (Lantz, 2013). This is based on finding the optimum hyperplane that separates the training data by the maximum margin (Wei and Chiu, 2015). SVM algorithms are powerful classifiers in cases where distribution is non-linear (Cortes and Vapnik, 1995). SVM models are tuned through thresholds in order to define residuals that consequently contribute to a more robust fit. Tuning parameters consist in *cost*, which is used to penalize models with large residuals, and a *loss* function that determines the degree of overfitting training models.

Partial Least Squares Discriminant Analysis (PLSDA). This algorithm classifies class through predictor combinations that optimally separate classes (Mevik and Wehrens, 2007). The tuning components of PLSDA are found through latent variables (components) that maximise classification accuracy. Once these components are identified, the PLSDA model is retrained to ensure accuracy in the final model. PLSDA in many cases has been identified as more powerful than standard Linear Discriminant Analysis (LDA) methods.

K-Nearest Neighbour (KNN). This unsupervised learning algorithm

Table 2
Comparison of different ML algorithm performance used for discerning carnivore agency (excluding wolves) through Yravedra et al.'s (2017) 2D 7-landmark model in Shape space.

Model	Model performance		95% confidence interval		Hyena		Jaguar		Lion	
	Kappa	Accuracy	Lower	Upper	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.
NNET	0.54	0.69	0.63	0.74	1	0.50	0.90	1	0.95	1
PLSDA	0.79	0.86	0.82	0.90	0.90	0.83	1	0.80	0.99	0.97
NB	0.82	0.88	0.84	0.91	0.98	0.77	0.88	0.90	0.97	0.96
CTREE	0.9	0.93	0.90	0.96	0.94	0.95	0.99	0.87	0.95	1
MDA	0.92	0.95	0.91	0.97	0.98	0.87	0.94	0.96	1	1
KNN	0.99	0.99	0.98	1	0.99	1	1	0.98	1	1
RF	1	1	0.99	1	1	1	1	1	1	1
C5.0	1	1	0.99	1	1	1	1	1	1	1
SVM	1	1	0.99	1	1	1	1	1	1	1

classifies unlabelled data by assigning them the class of the most similar labelled examples. This algorithm works well in samples with many variables and performs well when there are well-defined labelled sets. The main advantage of this algorithm is that it is easy to train. KNN identifies *k* cases in the sample as the nearest in similarity. Unlabelled cases are subsequently assigned by similarity. To predict the location of testing data in the predictor space, different *k* models are tested and compared to an error/accuracy parameter. To overcome the bias-variance trade-off an intermediate *k* value is usually selected. Larger *k* values tend to reduce the bias of variance but small patterns may go unnoticed (Domínguez-Rodrigo, 2018).

Mixture Discriminant Analysis (MDA). Initially conceived as an extension of LDA, MDA is built upon class-specific distributions combined into a single multivariate distribution. This is done by creating a per-class mixture (Kuhn and Johnson, 2013), consisting in the separation of class-specific means from class-specific covariance structures. MDA can be tuned using parameters in the number of distributions per class or sub-classes.

Naive Bayes (NB). NB algorithms apply Bayes' theorem through estimation of class probabilities and likelihoods on observed predictions. This results in dynamic estimates of posterior probabilities of classes. The conditional probability is then used for model classification. NB assumes that all predictors are independent and uses a nonparametric density modelling process.

Decision tress using the C5.0 algorithm (C5.0). This algorithm implements decision trees, employing ML techniques to ensure model accuracy. Unlike standard decision tress, the C5.0 algorithm tuning methods can reach comparable complexities to NN or SVM models. This model works through recursive partitioning of data. The C5.0 algorithm employs *k*-fold cross validation, dividing the sets into *k* subsets during the training and testing phases. Variance estimation decreases as *k* increases.

Conditional Inference Tress (CTREE). CTREEs are similar to standard decision tress in as much as they recursively partition data through univariate splits on dependent variables. The main difference presented in this algorithm is that CTREEs adapt significance test procedures in order to select the variables that maximise prediction accuracy (Wei and Chiu, 2015).

Random Forest (RF). RF is a more robust version of CTREE. Much like CTREE, the RF algorithm uses a small random number of the data set variables, instead of all the variables. Each selection produces an independent tree. The random variable selection is performed through bootstrap aggregation, known as bagging, splitting the data into multiple data sets for testing. These observations are referred to as out-of-bag (OOB) observations. RF produces estimates of how many iterations are needed to minimize the OOB error. After selecting a number of trees, the algorithm averages the results and produces a robust classification method. This avoids overfitting of results to data, as is more common in standard decision and regression trees.

More information on ML models used in taphonomy can be consulted in Domínguez-Rodrigo (2018).

2.5. Fine tuning GMM models

Common malpractice in GMM studies is the assumption that the first PC and CV scores are the most important for sample differentiation. Theoretically, the first few PC scores represent the highest percentage of sample variance and covariance, however, in some cases analytical scores with lower percentiles are useful to identify deviations from the majority (Albrecht, 1992). In the case of tooth scores, the first 10 PC scores were successful in most statistical applications, however the PC scores of Aramendi et al.'s (2017a) 3D 17-landmark model proved to be problematic due to the high number of PC scores generated through PCA (51 in total), and their small percentage of represented variance.

One of the advantages of ML techniques is the ability to evaluate model performance through examining the importance of each variable. In this case we used each model evaluation of variable importance to observe the classification power for each PC score. As expected, some PC scores are more important for classification than others. Reduction of the number of PC scores used can prune statistical noise while avoiding overfitting. Once the ML algorithms had identified the top performing PC scores for agent classification, the models were retuned using only these PC scores.

3. Results

3.1. ML differentiation of tooth scores

ML algorithms based on the information extracted from PC scores successfully construct powerful classification models (Tables S2 and S3). ML models in both shape and form space are relatively effective, though results tend to be more accurate after removing wolves from the analysis (Table 2). Additionally, score classification model rates are higher and more precise when using pure shape information, rather than when including centroid size in form space. At least 3 ML algorithms are able to reach 100% classification rates, with C5.0, RF and SVM being the best performing models for discerning carnivore agency via score cross-section morphology (Table 2). Nevertheless, sensitivity, specificity and accuracy values are high in most cases (Fig. 4), and ML performance is much more powerful than traditional GMM LDA models (Tables 3, 4, S4 and S5).

3.2. ML differentiation of tooth pits

Much like the case of tooth scores, 3D tooth pit models repeatedly produce up to 100% classification rates (Fig. 5), with SVM, RF, C5.0, KNN, MDA and PLSDA providing the highest degree of precision

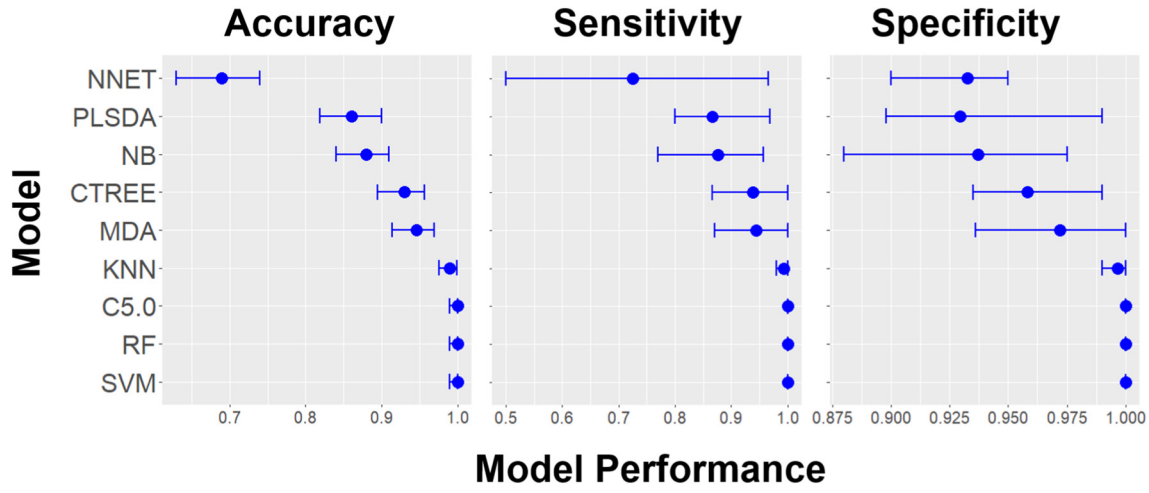


Fig. 4. Model performance data for machine learning algorithms when processing and classifying tooth score morphology.

Table 3
LDA performance comparing Lion, Hyena, Jaguar and Wolf tooth marks in both shape and form space for both GMM models. Specificity and sensitivity values are averaged, for full details consult Table S1.

	Performance		95% confidence interval			
	Kappa	Accuracy	Lower	Upper	Spec.	Sens.
2D shape	0.32	0.50	0.41	0.59	0.73	0.60
2D form	0.35	0.51	0.42	0.60	0.84	0.51
3D shape	0.13	0.35	0.25	0.46	0.78	0.34
3D form	0.16	0.37	0.27	0.48	0.79	0.37

Table 4
LDA performance comparing only Lion, Hyena and Jaguar tooth marks. Scores are studied in shape space while pits are studied in form space. Specificity and sensitivity values are averaged, for full details consult Table S2.

	Model performance		95% confidence interval			
	Kappa	Accuracy	Lower	Upper	Spec.	Sens.
Scores	0.46	0.64	0.54	0.73	0.81	0.48
Pits	0.14	0.43	0.31	0.56	0.70	0.52

(Table 5). But contrary to 2D cross-section profiles, higher classification models are generated when combining shape and centroid size in form space (Tables S6 and S7).

The analysis including all 4 carnivore species (spotted hyenas, lions, jaguars and wolves) based on Aramendi et al.'s (2017a) model provides relatively low classification results (Tables S6 and S7). For this study, 4 consequent tuning tests were carried out on PC scores using ML algorithms. The first test includes all PC scores; the second model tuning includes the top 10 PC scores identified through test 1; the third tuning takes in the top 15 PC scores; and the final tuning procedure includes the top 20 PC scores. All tables can be consulted in Table S8. Model performance was then re-evaluated for each testing cycle (Fig. 6). According to the results, the inclusion of at least the 15 first PC scores obtains the best performance accuracy for discerning carnivore agency through tooth pit morphology.

Nevertheless, here again, the exclusion of wolf tooth pits from the analysis generates a much more powerful model without the need for PC tuning (Table 5, Fig. 6).

4. Discussion

Initial GMM studies of both scores and pits originally produced significant differences between most carnivore species (Arriaza et al., 2017; Yravedra et al., 2017; Aramendi et al., 2017a). Such differences are reflected in significant p values ($p < 0.05$) obtained via MANOVA

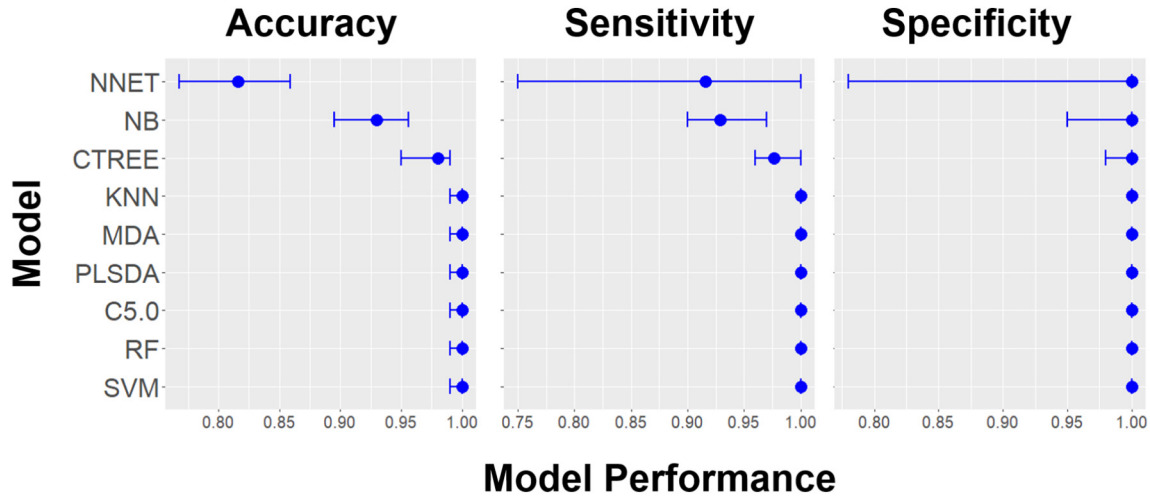


Fig. 5. Model performance data for machine learning algorithms when processing and classifying tooth pit morphology.

Table 5
Comparison of different ML algorithm performance used for discerning carnivore agency (excluding wolves) through Aramendi et al. (2017a) 3D 17-landmark model in Form space.

Model	Model performance		95% confidence interval		Hyena		Jaguar		Lion	
	Kappa	Accuracy	Lower	Upper	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.
NNET	0.73	0.82	0.77	0.86	1	0.75	0.78	1	0.82	1
NB	0.90	0.93	0.90	0.96	0.95	0.90	1	0.92	0.95	0.97
CTREE	0.97	0.98	0.95	0.99	1	1	0.98	0.96	0.98	0.97
PLSDA	1	1	0.99	1	1	1	1	1	1	1
MDA	1	1	0.99	1	1	1	1	1	1	1
KNN	1	1	0.99	1	1	1	1	1	1	1
RF	1	1	0.99	1	1	1	1	1	1	1
C5.0	1	1	0.99	1	1	1	1	1	1	1
SVM	1	1	0.99	1	1	1	1	1	1	1

tests and a relatively clear separation of groups in the graphs and the distances computed after Canonical Variance Analysis (CVA). Regardless, in these original studies, PCAs still presented significant overlapping in the majority of cases. LDA results were usually able to achieve relatively high classification rates (Aramendi et al., 2017a; Arriaza et al., 2017; Yravedra et al., 2017), however the efficiency and performance of LDA as classification method is not as powerful as desirable (Tables 3 and 4). Nevertheless, the exclusion of wolves from the analysis appears to clear up some confusion (Table 2), as previously observed in other studies (Yravedra et al., 2018). In sum, GMM results *alone* show that the 2D analysis of tooth scores produce better results when performed in shape space, while the 3D analysis of tooth pits is more accurate when using form data (Fig. 7).

In this study a combined use of ML algorithms and taphonomic GMM models has been able to provide up to 100% classification when discerning between carnivore agencies. NNET and NB are generally the worst performing models, while SVM, C5.0 and RF consistently perform much better reaching exceptional results (Kappa = 1, Accuracy = 1).

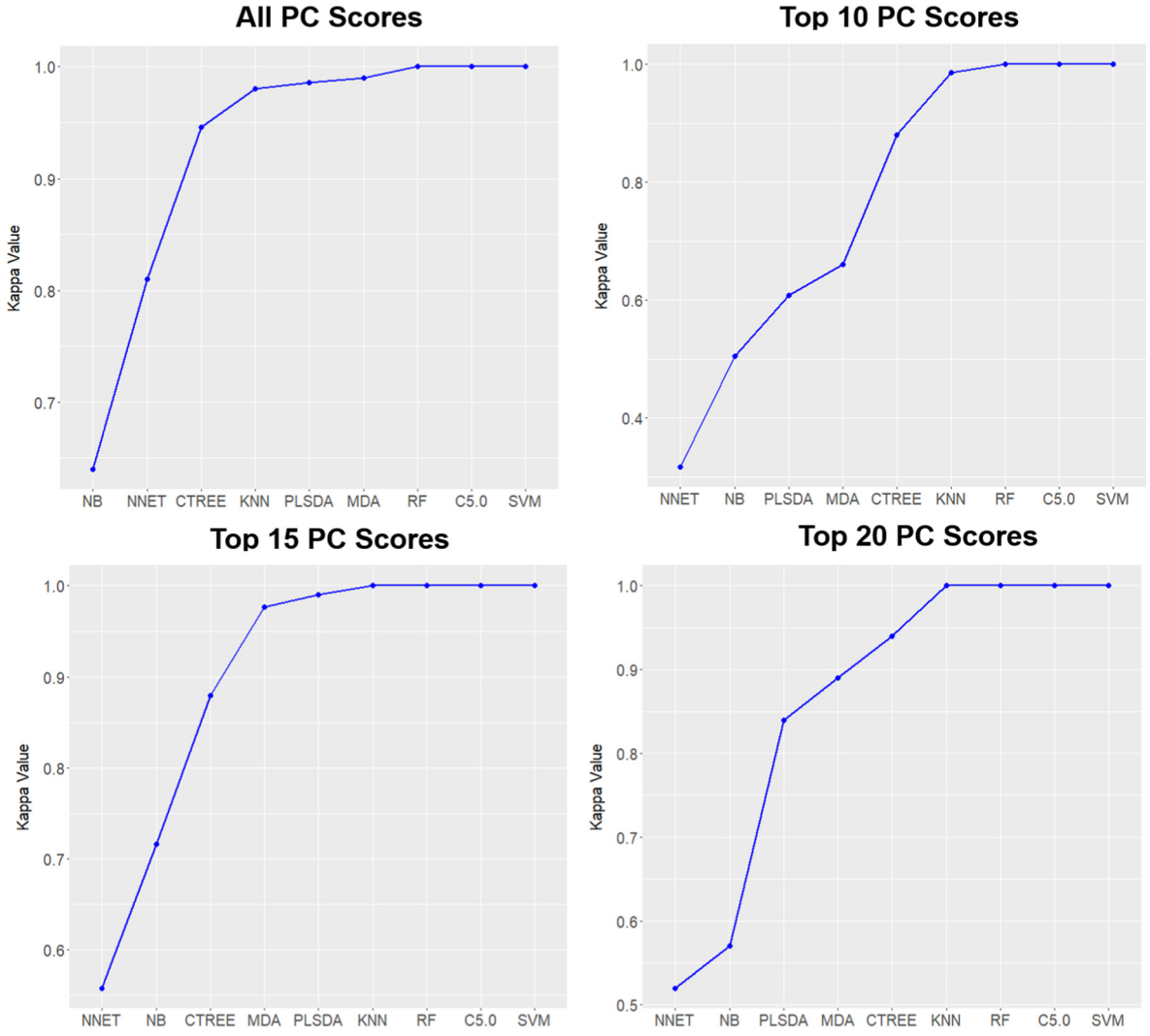


Fig. 6. Kappa value curves for each machine learning algorithm after each stage of the PC score tuning of Aramendi et al. (2017a) 3D seventeen landmark model.

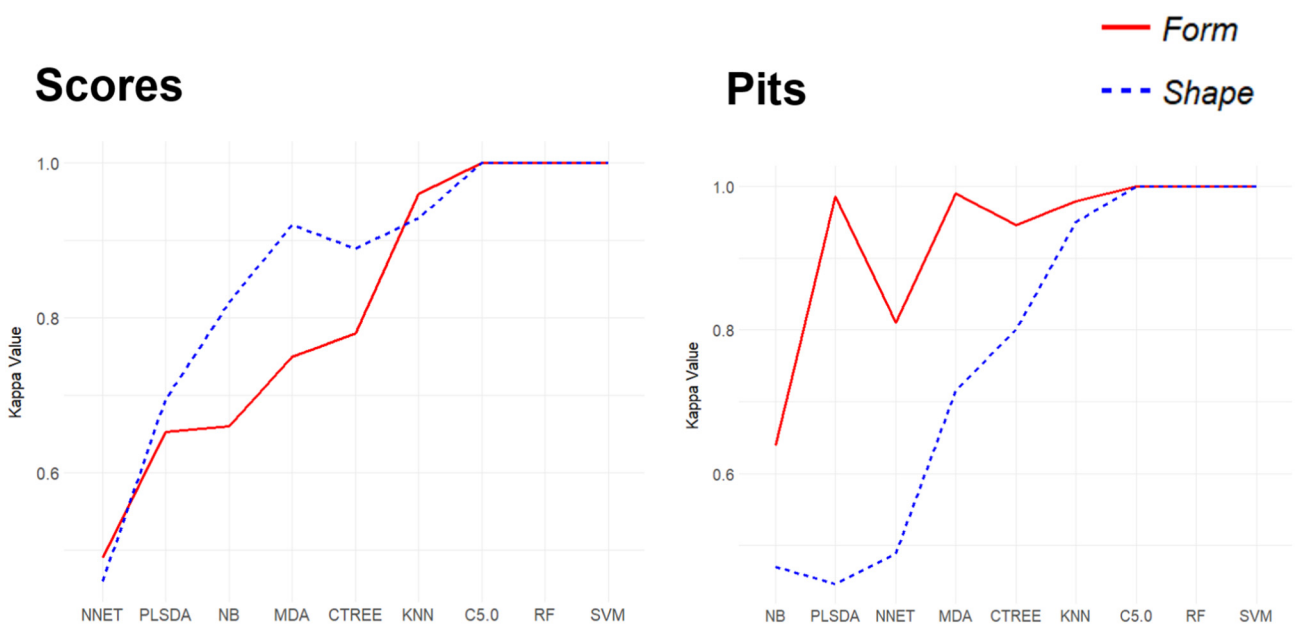


Fig. 7. Kappa value curves for machine learning algorithm performance when processing both pits and scores in shape space and form space.

CTREE, PLSDA, MDA and KNN are inconsistent, only providing high classification rates in some cases. As seen in the tuning of the 3D 17-landmark model, however, future use of ML and GMM based methods in classifying unknown taphonomic traces should take into account the possible effect that PC score selection could have on the quality of ML based classifications. Some algorithms, for example, were unable to provide 100% classification prior to tuning. The process of tuning, however, was able to highlight the potential of ML based methods on multiple accounts.

The poor performance of some models can be argued by the complexity of their internal architecture. In the case of NNET, if a more developed model is constructed using optimised hyperparameters, such as those typical of Deep Learning (Chollet and Alaire, 2017), these refined Deep Neural Networks are more likely to produce higher classification results. The optimisation of a Deep Learning model that confronts GMM data, therefore, would be a useful development for the future.

In general, ML based differentiation of carnivore agency through Aramendi et al.'s (2017a) 3D 17-landmark model provides the best results (Fig. 8). Nevertheless, Yravedra et al.'s (2017) 2D 7 landmark model is still capable of producing powerful classification models. Most models achieve Kappa values above the margin of 0.8, thus, they can be considered powerful predictive models (Lantz, 2013). According to our results, tooth scores should be preferably analyzed on shape data by means of at least RF, C5.0 and SVM algorithms. In the case of tooth pits, the analysis should be rather performed in form space, however, care should be taken when dealing with a large number of PC scores. Tuning of models is highly recommended when dealing with great overlapping samples in PCAs.

These new statistical applications may be able to provide a deeper understanding of hominin-carnivore interactions, as considered to be a critical component in human evolutionary studies. The importance of understanding hominin-carnivore reactions is two-fold. Firstly, because carnivores may be responsible for bone accumulations and taphonomists must be able to differentiate between anthropogenic and carnivore-accumulated bone assemblages. Secondly, because some key

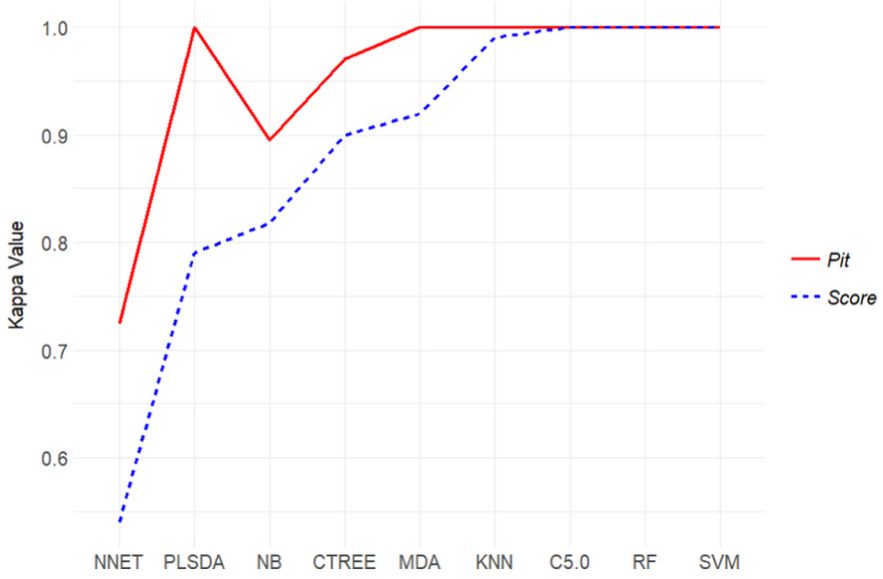


Fig. 8. Final kappa value curves for machine learning algorithm performance when processing pits in form space and scores in shape space.

paleoanthropological sites have been considered palimpsests in which both humans and carnivores may have played an important role in bone modification, such as those from Olduvai Gorge (Domínguez-Rodrigo et al., 2007). The improvement of classical BSM variables is therefore of great importance to help discerning the accumulating agents of bone assemblages at archaeological sites.

From this perspective, the application of GMM and micro-photogrammetric techniques in studying BSMs have helped to avoid some shortcomings of previous taphonomic variables, such as the equifinality arisen from the study of tooth mark frequencies and dimensions (Blumenschine, 1988; Selvaggio and Wilder, 2001; Domínguez-Rodrigo and Piqueras, 2003; Andrés et al., 2012). While bidimensional studies of tooth marks have been successful in distinguishing different sized carnivores, within these groups of large and small animals, little differentiation is possible (Andrés et al., 2012). This equifinality is hard to overcome relying on these variables alone, and their sole use is not recommended. Combined with tooth mark frequencies, skeletal profiles, fracture patterns and taphotypes, better results have been achieved (Domínguez-Rodrigo et al., 2007; Pineda et al., 2015; Aramendi et al., 2017b; Saladié et al., 2017). This highlights the need for a combined use of multiple criteria when analysing an assemblage.

Recent applications of ML based techniques in the study of modern carnivore sites have efficiently addressed some of these questions (Arriaza and Domínguez-Rodrigo, 2016), considering multiple traditional variables. Alongside these approaches, very recently these algorithms have been applied to multiple BSMs qualitative criteria (Domínguez-Rodrigo, 2018; Domínguez-Rodrigo and Baquedano, 2018)

However, the introduction of ML in archaeology occurred much later than in other sciences (Michalski et al., 1983). In other fields, the development of ML algorithms as part of the field of artificial intelligence have provided a much more efficient means of handling large amounts of raw data, using a more structured means of processing these data sets (Lantz, 2013; Paterson and Gibson, 2017). An additional advantage presented by ML is its versatility, using tuning functions to reach higher degrees of accuracy based on performance (Wolpert, 1996). With the inclusion of Computer Vision and Deep Learning techniques (Chollet and Alaire, 2017; Paterson and Gibson, 2017), the possibilities artificial intelligence may provide for future archaeological studies are ever-growing.

A certain degree of caution is required when assuming ML is the absolute solution to eliminating equifinality, however. A problematic concept that requires further confrontation lies in data collection, and the objectivity behind obtaining such information. The development of GMM over the years has come to provide increasingly more objective means of processing features of shape and form, however the most reliable and efficient means of obtaining the 3D models used is still under investigation (Maté-González et al., 2017b, c; Courtenay et al., 2017, Under review). Likewise, comparisons between different landmark models are still pending further investigation. Nevertheless, as presented here, combining both GMM and ML approaches is still able to provide interesting results. From this perspective the benefits of a hybrid ML and GMM approach to taphonomy could produce an important development in BSM analysis.

5. Conclusions

This paper presents the first combination of advanced statistical analysis using GMM and ML methods to discern carnivore agency through different types of tooth marks. The future of ML and GMM based methods is promising. Our results provide the first combined inclusion of both statistical techniques, achieving 100% classification rates in multiple models. So far, this technique has only been applied to determine carnivore agency, but can be expanded to include other GMM models.

The possibility provided by the combination of ML and GMM based techniques could be of utmost relevance when applied to Lower Pleistocene sites, where the interaction of several modifying agents is still subject of discussion. Though the comparative sample still needs to be increased, adding more carnivore species and tooth marks, the methodological approach described throughout this paper opens up the door to new work lines in taphonomy. Furthermore, the development of this line of research could also shed light into certain debates regarding early human behaviour.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.palaeo.2019.03.007.

Acknowledgements

Firstly, we would like to thank all the staff and members of both the IPHES and the Rovira i Virgili University. Along these lines, we would also like to acknowledge the staff at the Parque de la Naturaleza de Cabárceno and Santiago Borragan for providing the samples used in our experiments. Finally we would like to thank Abel Moclán for his advice and comments during the process of writing this paper.

Additionally, we would like to thank the TIDOP Group from the Department of Cartographic and Land Engineering of the Higher Polytechnics School of Avila, University of Salamanca, for the use of their tools and facilities. We want to recognize the technical support provided by C.A.I. Arqueometry and Archaeological Analysis from Complutense University which has been very useful in carrying out the present paper. J.A. would like to thank Fundación La Caixa and the Spanish Education, Culture and Sports Ministry (FPU15/04585) for funding her postgraduate education program. The support of the DST-NRF Centre of Excellence in Palaeosciences (CoE-Pal) towards this research is also acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the CoE. Finally, this work was developed within the general framework of the Spanish MINECO-FEDER project CGL2015-65387-C3-1-P, the Catalan AGAUR project 2017-SGR-1040, and the URV project 2017-PFR-URV-B2-91.

References

Albrecht, G.H., 1992. Assessing the affinities of fossils using canonical variates and generalized distances. Hum. Evol. 7 (4), 49–69.

Alcántara, V.G., Egido, R.B., del Pino, J.M.B., Ruiz, A.B.C., Vidal, A.I.E., Aparicio, Á.F., Calleja, S.H., Jiménez, A.I., González, M.M., Gil, M.P., Tello, V.P., Calvo, J.R., Yravedra, J., Vidal, A.S., Domínguez-Rodrigo, M., 2006. Determinación de procesos de fractura sobre huesos frescos: un sistema de análisis de los ángulos de los planos de fracturación como discriminador de agentes bióticos. Trab. Prehist. 63 (1), 37–45.

Andrés, M., Gidna, A.O., Yravedra, J., Domínguez-Rodrigo, J., 2012. A Study of Dimensional differences of tooth marks (pits and scores) on bones modified by small and large carnivores. Journal of Archaeological Anthropological Sciences. 4, 209–219.

Aramendi, J., Maté-González, M.Á., Yravedra, J., Ortega, M.C., Arriaza, M.C., González-Aguilera, D., Baquedano, E., Domínguez-Rodrigo, M., 2017a. Discerning carnivore agency through the three-dimensional study of tooth pits: Revisiting crocodile feeding behaviour at FLK-Zinj and FLK NN3 (Olduvai Gorge, Tanzania). Palaeogeogr. Palaeoclimatol. Palaeoecol. 488, 93–102.

Aramendi, J., Uribelarrea, D., Arriaza, M.C., Arráiz, H., Barboni, D., Yravedra, J., Cruz Ortega, M., Gidna, A., Mabulla, A., Baquedano, E., Domínguez-Rodrigo, M., 2017b. The Paleoecology and Taphonomy of AMK (Bed I, Olduvai Gorge) and its Contributions to the Understanding of the "Zinj" Paleolandscape, Paleogeography, Paleoclimatology. Paleoecology. 488, 35–49.

Arriaza, M.C., Domínguez-Rodrigo, M., 2016. When Felids and Hominins Ruled at Olduvai Gorge: a Machine Learning Analysis of the Skeletal Profiles of the Non-Anthropogenic Bed I Sites. Quat. Sci. Rev. 139, 43–52.

Arriaza, M.C., Domínguez-Rodrigo, M., Yravedra, J., Baqudano, E., 2016. Lions as bone accumulators? Paleontological and ecological implications of a modern bone assemblage from Olduvai Gorge. PLoS One 11 (5), e0153797.

Arriaza, M.C., Yravedra, J., Domínguez-Rodrigo, M., Mate-González, M.Á., Vargas, E.G., Palomeque-González, J.F., Aramendi, J., González-Aguilera, D., Baquedano, E., 2017. On applications of micro-photogrammetry and geometric morphometrics to studies of tooth mark morphology: the modern Olduvai Carnivore Site (Tanzania). Palaeogeogr. Palaeoclimatol. Palaeoecol. 488, 103–112.

Arriaza, M.C., Aramendi, J., Maté-González, M.Á., Yravedra, J., Baquedano, E., González-Aguiler, D., Domínguez-Rodrigo, M., 2018. Geometric morphometric analysis of tooth pits and identification of felid and hyenid agency in Bone modification. Quat. Int. https://doi.org/10.1016/j.quaint.2018.11.023.

Bartram, L.E., Marean, C.W., 1999. Explaining the "Klasies Pattern": Kua

Ethnoarchaeology, the Die Kelders Middle Stone Age Archaeofauna. long bone fragmentation and carnivore ravagin, Journal of Archaeological Science. 26, 9–29.
Binford, L.R., 1981. Bones: Ancient Men and Modern Myths. Academic Press.
Blumenschine, R., 1988. An experimental model of the timing of hominid and carnivore influence on archaeological bone assemblages. J. Archaeol. Sci. 15, 483–502.
Blumenschine, R., 1995. Percussion marks, tooth marks and experimental determinations of the timing of hominid and carnivore access to long bones at FLK Zinjanthropus, Olduvai Gorge, Tanzania. J. Hum. Evol. 29 (1), 21–51.
Blumenschine, R., Marean, C.W., 1993. Long bone marrow yields of Some African ungulates. J. Archaeol. Sci. 20, 555–587.
Bookstein, F., 1991. Morphometric Tools for Landmark Data: Geometry and Biology. In: New York. Press, Cambridge University.
Brain, C.K., 1967. Hottentot food remains and their bearing on the interpretation of fossil bone assemblages. Scientific Papers of the Namib Desert Research Station 32, 1–11.
Brain, C.K., 1969. The contribution of Namib Desert Hottentots to an understanding of Australopithecine bone accumulations. Scientific Papers of the Namib Desert Research Station 39, 13–22.
Brain, C.K., 1981. The hunter or the hunted. An Introduction to African Cave Taphonomy. The University of Chicago press, Chicago.
Bunn, H.T., 1981. Archaeological evidence for meat eating by Plio-Pleistocene hominids from Koobi Fora and Olduvai Gorge. Nature. 291, 574–577.
Bunn, H.T., Pickering, T.R., 2010. Methodological recommendations for ungulate mortality analyses in paleoanthropology. Quat. Res. 74, 388–394.
Capaldo, S., 1997. Experimental determinations of carcass processing by plio-pleistocene hominids and carnivores at FLK 22 (Zinjanthropus. Olduvai Gorge, Tanzania). J. Hum. Evol. 33, 555–597.
Capaldo, S.D., Blumenschine, R.J., 1994. A quantitative diagnosis of notches made by hammerstone percussion and carnivore gnawing in bovid long bones. Am. Antiq. 59, 724–748.
Chollet, F., Allaire, J.J., 2017. Deep Learning with R. Manning Publications, New York.
Core-Team, 2015. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.Rproject.org/, 2015. Accessed the 14th of February, 2018.
Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. 20, 273–297.
Courtenay, L.A., Yravedra, J., Maté-González, M.Á., Aramendi, J., González-Aguilera, D., 2017. 3D analysis of cut marks using a new geometric morphometric methodological approach. Journal of Archaeological and Anthropological Sciences. https://doi.org/10.1007/s12520-017-0554-x.
Courtenay, L.A., Yravedra, J., Huguet, R., Ollé, A., Aramendi, J., Maté-González, M.Á., González-Aguilera, D., 2019. New taphonomic advances in 3D digital microscopy: a morphological characterisation of trampling marks. Quat. Int. https://doi.org/10.1016/j.quaint.2018.12.019. (under review).
Cruz-Uribe, K., 1991. Distinguishing hyena from hominid bone accumulations. J. Field Archaeol. 18 (4), 467–486.
Dart, R.A., 1957. The Osteodontokeratic culture of Australopithecus africanus. Transvaal Museum Memoires 10, 1–105.
Dawkins, W.B., 1863. Wookey Hole hyaena den. Procedings of the Somersetshire Archaeology and Natural History Society. 11 (2), 197–219.
Delaney-Rivera, C., Plummer, T.W., Hodgson, J.A., Forrest, F., Hertel, F., Oliver, J.S., 2009. Pits and pitfalls: taxonomic variability and patterning in tooth mark dimensions. J. Archaeol. Sci. 36, 2597–2608.
Domínguez-Rodrigo, M., 1999. Flesh availability and bone modifications in carcasses consumed by lions: Paléoecologucal relevance in hominid foraging patterns. Palao 2189. Palaeogeography, Palaeoclimatology. Plaeoecology. 149, 373–388.
Domínguez-Rodrigo, M., 2001. A study of carnivore competition in riparian and open habitats of modern Savannas and its implications for hominid behavioral modelling. J. Hum. Evol. 40, 77–98.
Domínguez-Rodrigo, M., 2002. Hunting and scavenging by early humans: the state of the debate. J. World Prehist. 16, 1–54.
Domínguez-Rodrigo, M., 2018. Successful Classification of Experimental Bone Surface Modifications (BSM) through Machine Learning. A Solution to the Controversial Use of BSM in Paleoanthropology, Archaeological and Anthropological Sciences, Algorithms. https://doi.org/10.1007/s12520-018-0684-9.
Domínguez-Rodrigo, M., Baquedano, E., 2018. Distinguishing butchery cut marks from crocodile bite marks through machine learning methods. Sci. Rep. https://doi.org/10.1038/s41598-018-24071-1.
Domínguez-Rodrigo, M., Pickering, T.R., 2010. Un Estudio Tafonómico Multivariante de las Acumulaciones de Fauna de Hiénidos (Crocuta crocuta) y Félidos (Panthera pardus). 1º Reunión de Científicos Sobre Cubiles de Hiena (y otros grandes carnivoros) en los Yacimientos Arqueológicos de la Peninsula Ibérica 45–60.
Domínguez-Rodrigo, M., Piqueras, A., 2003. The use of tooth pits to identify carnivore taxa in tooth-marked Archaeofaunas and their relevance to reconstruct hominid carcass processing behaviours. J. Archaeol. Sci. 30 (11), 1385–1391.
Domínguez-Rodrigo, M., Barba, R., Egeland, C.P., 2007. Deconstructing Olduvai. Springer, The Netherlands.
Domínguez-Rodrigo, M., Gidna, A.O., Yravedra, J., Musiba, C., 2012. A comparative neotaphonomic study of felids, hyaenids and canids: an analogical framework based on long bone modification patterns. Journal of Taphonomy. 10 (3), 147–164.
Domínguez-Rodrigo, M., Yravedra, J., Oganista, E., Gidna, A., Fourvel, J.B., Baquedano, E., 2015. A new methodology approach to the taphonomic study of paleontological and archaeological faunal assemblages: a preliminary case study from Olduvai Gorge (Tanzania). J. Archaeol. Sci. 59, 35–53.
Dryden, I.L., Mardia, K.V., 1998. Statistical Shape Analysis. John Wiley & Sons, Chichester.
Egeland, A.G., Egeland, C.P., Bunn, H.T., 2008. Taphonomic analysis of a modern spotted hyena (*Crocuta crocuta*) den from Nairobi, Kenya. Journal of Taphonomy. 6 (3–4), 275–299.
Faith, J.T., Marean, C.W., Behrensmeyer, A.K., 2007. Carnivore competition, bone destruction, and bone density. J. Archaeol. Sci. 34 (12), 2025–2034.
Fawcett, T., 2006. An Introduction to ROC analysis. Pattern Recogn. Lett. 27, 861–874.
Galán, A.B., Rodríguez, M., de Juana, S., Domínguez-Rodrigo, M., 2009. A new experimental study on percussion marks and notches and their bearing on the interpretation of hammerstone-broken faunal assemblages. J. Archaeol. Sci. 36, 776–784.
Gidna, A., Yravedra, J., Domínguez-Rodrigo, M., 2013. A cautionary note on the use of captive carnivores to model wild predator behavior: a comparison of bone modification patterns on long bones by captive and wild lions. J. Archaeol. Sci. 40, 1903–1910.
Gidna, A.O., Kisui, B., Mabulla, A., Musiba, C., Domínguez-Rodrigo, M., 2014. An ecological neo-taphonomic study of carcass consumption by lions in Tarangire National Park (Tanzania) and its relevance for human evolutionary biology. Quat. Int. 322–323, 167–180.
González-Aguilera, D., López Fernández, L., Rodríguez-González, P., Guerrero, D., Hernandez-Lopez, D., Remondino, F., Menna, F., Nocerino, E., Toschi, I., Ballabeni, A., Gaiani, M., 2016. Development of an all-purpose free photogrammetric tool. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Conference.
Günther, F., Fritsch, S., 2010. neuralnet. training of neural networks, The R Journal. 2 (1), 30–38.
Haynes, G., 1982. Utilization and Skeletal Disturbances of North American Prey Carcasses. Artic 35, 266–281.
Haynes, G., 1983a. Frequencies of spiral and green-bone fractures on ungulate limb bones in modern surface assemblages. Am. Antiq. 48, 102–114.
Haynes, G., 1983b. A guide for differentiating mammalian carnivore taxa responsible for gnaw damage to herbivore limb bones. Paleobiology. 9, 164–172.
Houtson D.C., 1979. The adaptations of scavengers in A. R. E. Sinclair and M. Narton Griffiths eds 263-286
Kerbis-Peterhans, J.C., 1990. The Role of Porcupines, Leopards and Hyaenas in Ungulate Carcass Dispersal: Implications for Paleoanthropology. University of Chicago, Chicago, Department of Anthropology.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, London.
Lantz, B., 2013. Machine Learning with R. Packt Publishing Ltd., Birmingham.
Lyman, R.L., 1994. Vertebrate Taphonomy. Cambridge University Press.
Marean, C.W., Frey, C.J., 1997. Animal bones from caves to cities: reverse utility curves as methodological artifacts. Am. Antiq. 62 (4), 698–711.
Martín, F.M., 2008. Bone crunching felids at the end of the Pleistocene in Fuego-Patagonia, Chile. Journal of Taphonomy 6 (3–4), 337–372.
Maté-González, M.Á., Yravedra, J., González-Aguilera, D., Palomeque-González, J.F., Domínguez-Rodrigo, M., 2015. Microphotogrammetric characterization of cut marks on bones. J. Archaeol. Sci. 62, 128–142.
Maté-González, M.Á., Palomeque-González, J.F., Yravedra, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2016. Micro-photogrammetric and morphometric differentiation of cut marks on bones using metal knives, quartzite and flint flakes. Journal of Archaeological and Anthropological Science. https://doi.org/10.1007/s12520-016-0401-5.
Maté-González, M.Á., Yravedra, J., Martín-Perea, D., Palomeque-González, J., San-Juan-Blazquez, M., Estaca-Gómez, V., Uribelarrea, D., Álvarez-Alonso, D., Cuartero, F., González-Aguilera, D., Domínguez-Rodrigo, M., 2017a. Flint and quartzite: distinguishing raw material through bone cut marks. Archaeometry. https://doi.org/10.1111/arcm.12327.
Maté-González, M.A., Aramendi, J., Yravedra, J., Blasco, R., Rosell, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2017b. Assessment of statistical agreement of three techniques for the study of cut marks: 3D Digital Microscope, Laser Scanning Confocal Microscopy and Micro-Photogrammetry. J. Microsc. 267 (3), 356–370. https://doi.org/10.1111/jmi.12575.
Maté-González, M.A., Aramendi, J., Yravedra, J., González-Aguilera, D., 2017c. Statistical comparison between low-cost methods for 3D characterization of cut-marks on bones. Remote Sens. https://doi.org/10.3390/rs9090873.
Mevik, B.H., Wehrens, R., 2007. The pls package: principal component and partial least squares regression in R. J. Stat. Softw. 18 (2), 1–23.
Michalski, R.S., Carbonell, J.G., Mitchell, T.M., 1983. Machine Learning: An Artificial Intelligence Approach, Volume I. Morgan Kaufman Publishers Inc., USA.
Mitteroecker, P., Gunz, P., 2009. Advances in geometric morphometrics. Evol. Biol. 36, 235–247.
Moclán, A., Domínguez-Rodrigo, M., 2018. An experimental study of the patterned nature of anthropogenic bone breakage and its impact on bone surface modification frequencies. J. Archaeol. Sci. 96, 1–13.
Moclán, A., Domínguez-Rodrigo, M., 2019. Classifying Bone Breakage Patterns: an Experimental Analysis of Fracture Planes to Discern Between Hominin and Carnivore Activity using Machine Learning (ML) Algorithms (accepted).
Paterson, J., Gibson, A., 2017. Deep Learning: A Practitioner's Approach. O'Reilly Media Inc, California.
Pickering, T.R., 2002. Reconsideration of Criteria for Differentiating Faunal Assemblages Accumulated by Hyenas and Hominids. Int. J. Osteoarchaeol. 12, 127–141.
Pickering, T.R., Bunn, H.T., 2007. The endurance running hypothesis and hunting and scavenging in Savanna Woodlands. J. Hum. Evol. 53, 434–438.
Pickering, T.R., Egeland, C.P., 2006. Experimental patterns of hammerstone percussion damage on bones: implications for inferences of carcass processing by humans. J. Archaeol. Sci. 33, 459–469.
Pickering, T.R., Marena, C.W., Domínuez-Rodrigo, M., 2003. Importance of Limb Bone Shaft fragments in zooarchaeology: a response to “on in situ attrition and vertebrate body part profiles” (2002), by M. C. Stiner, Journal of Archaeological Science. 30 (11), 1469–1482.

Pickering, T.R., Domínguez-Rodrigo, M., Egeland, C.P., Brain, C.K., 2005. The contribution of limb bone fracture patterns to reconstructing early hominid behaviour at Swartkrans cave (South Africa): archaeological application of a new analytical method. Int. J. Osteoarchaeol. 15, 247–260.

Pickering, T.R., Heaton, J.L., Zwodeski, S.E., Kuman, K., 2011. Taphonomy of bones from baboons killed and eaten by wild leopards in Mapungubwe National Park, South Africa. Journal of Taphonomy 9 (2), 117–159.

Pineda, A., Saladié, P., Huguer, R., Cáceres, I., Rosas, A., García-Tabernero, A., Estalrrich, A., Mosquera, M., Ollé, A., Vallverdú, J., 2015. Coexistence among large predators during the Lower Paleolithic at the Site of La Mina (Barranc de la Boella, Tarragona, Spain). Quat. Int. 388, 177–187.

Rohlf, F.J., 1999. Shape statistics: procrustes superimpositions and tangent spaces. J. Classif. 16 (2), 197–223.

Ruiter, D.J., Berger, L.R., 2000. Leopards as taphonomic agents in dolomitic caves – implications for bone accumulations in the hominid-bearing deposits of South Africa. J. Archaeol. Sci. 27 (8), 665–684.

Saladié, P., Fernández, P., Rodríguez-Hidalgo, A., Huguet, R., Pineda, A., Cáceres, I., Marín, J., Vallverdú, J., Carbonell, E., 2017. The TD63.3 Faunal Assemblage of the Gran Dolina site (Atapuerca, Spain): a Late Early Pleistocene Hyena Den. DOI: https://doi.org/10.1080/08912963.2017.1384476.

Schaller, G.B., Lowther, G.R., 1969. The relevance of carnivore behavior to the study of early hominds. South W. Journal of Anthropology 25, 307–341.

Selvaggio, M.M., 1994. Identifying the Timing and Sequence of Hominid and Carnivore Involvement with Plio-Pleistocene Bone Assemblages from Carnivore Tooth Marks and Stone Tool Butchery Marks on Bone Surfaces. PhD Thesis. New Brunswick: Rutgers University.

Selvaggio, M.M., Wilder, J., 2001. Identifying the involvement of multiple carnivore taxon with archaeological bone assemblages. J. Archaeol. Sci. 28, 465–470.

Slice, D.E., 2001. Landmark coordinates aligned by procrustes analysis do not lie in Kendall's shape space. Syst. Biol. 50 (1), 141–149.

Stanford, C.B., Bunn, H.T., 2001. Meat Eating and Human Evolution. Oxford University Press, New York.

Stiner, M., 1994. Honor Among Thieves: A Zooarvheological Study of Neandertal Ecology. Princeton University press, Princeton.

Stiner, M.C., 2004. Comparative ecology and taphonomy of spotted hyenas, humans, and wolves in Pleistocene Italy. Rev. Paleobiologie 23, 771–785.

Sutcliffe, A.J., 1970. Spotted hyena: Crusher, gnawer, digester and collector of bones. Nature 227, 110–1113.

Wei, Y., Chiu, D., 2015. Machine Learning with R Cookbook. Packt Publishing, Birmingham.

Willey, P., Snyder, L.M., 1989. Canid modification of human remains: implications for time-since-death estimations. J. Forensic Sci. 34, 894–901.

Wolpert, D.H., 1996. The existence of a priori distinctions between learning algorithms. Neural Comput. 8, 1391–1420.

Yravedra, J., Lagos, L., Bárcena, F., 2011. A taphonomic study of wild wolf (Canis lupus) modification of horse bones in northwestern Spain. Journal of Taphonomy. 9 (1), 37–65.

Yravedra, J., Lagos, L., Bárcena, F., 2012. The wild wolf (Canis lupus) as a dispersal agent of animal carcasses in Northwestern Spain. J. Taphon. 10, 219–238.

Yravedra, J., García-Vargas, E., Maté-González, M.Á., Aramendi, J., Palomeque-González, J.F., Vallés-Iriso, J., Matesanz-Vicente, J., González-Aguilera, D., Domínguez-Rodrigo, M., 2017. The use of micro-photogrammetry and geometric morphometrics for identifying carnivore agency in bone assemblages. J. Archaeol. Sci. Rep. 14, 106–115.

Yravedra, J., Aramendi, J., Maté-González, M.Á., Courtenay, L.A., González-Aguilera, D., 2018. Differentiating percussion pits and carnivore tooth pits using 3D reconstruction and geometric morphometrics. PLoS One. https://doi.org/10.1371/journal.pone.0194324.