1 2 9 0

UNIVERSIDADE Ð
COIMBRA

Francisca Isabel Oliveira Fernandes

CLIMATE CHANGE AND HOMININ ORIGINS:
AN ECOMORPHOLOGICAL AND ECOMETRIC ANALYSIS IN BOVIDS TO RECONSTRUCT THE AFRICAN EARLY PLIOCENE ENVIRONMENT

Dissertação no âmbito do Mestrado em Evolução e Biologia Humanas orientada pela Professora Doutora Cláudia Isabel Soares Umbelino e pelo Doutor Thomas A. Püschel e apresentada ao Departamento de Ciências da Vida, da Faculdade de Ciências e Tecnologia.

outubro de 2021

Faculdade de Ciências e Tecnologia
Departamento de Ciências da Vida

Francisca Isabel Oliveira Fernandes

Climate Change and hominin origins: An Ecomorphological and Ecometric Analysis in Bovids to Reconstruct the African Early Pliocene Environment

VOLUME 1

Dissertação no âmbito do Mestrado em Evolução e Biologia Humanas orientada pela Professora Doutora Cláudia Isabel Soares Umbelino e pelo Doutor Thomas A. Püschel e apresentada ao Departamento de Ciências da Vida, da Faculdade de Ciências e Tecnologia.

outubro de 2021

1 2 9 0
UNIVERSIDADE DE
COIMBRA

This page is intentionally left blank.

Table of Contents

List of Figures

List of Tables

Abstract

The influence of climate change on human evolution is much debated. Nevertheless, hominin fossils are scarce and different research directions, such as taxon-free methods, attempt to confront this reality. This project aims to apply these methods, using various machine learning (ML) techniques, to reconstruct the palaeoenvironmental aspects of early Pliocene African sites related to *Australopithecus anamensis*.

The proposed study used published data from astragalar measurements of bovids, taken from Kanapoi (n=31) and Allia Bay (n=15) fossils, and from extant individuals (n= 187). Two analyses were performed, ecomorphology where individuals from the extant sample were classified into different habitats and soil moist categories, and ecometrics, where the extant sample as a community was used to predict various precipitation and temperature variables, as well as classify land cover categories. Finally, several ML algorithms were trained for each analysis using data from the extant taxa to infer the possible environment of the fossil sample.

The obtained results reveal better classification and prediction rates when applying different algorithms to the analyses in comparison to linear discriminant analysis (LDA), the method commonly used in these analyses. Furthermore, the ecomorphological analyses show mostly intermediate habitats and a mix of dry and wet categories for both sites. In a similar way, ecometric analyses show sites with tropical temperatures (23.8 and 25.7ºC), humid (>850 mm) but highly variable and consequently highly heterogeneous vegetation, combining evergreen broadleaf forests and herbaceous woodlands, but also more open systems such as grasslands. These two sites fall within a highly complex and variable context that can contribute to sudden changes in the environment. This seems to indicate that *A. anamensis* had thrived in these types of environments.

Ultimately, ecometric methods seem more promising, as their results allow a clearer attribution to a potential climate. Hence, it would be interesting to extend the applied approach to other important sites associated with our own evolutionary history.

Keywords

Palaeoenvironments; Kanapoi; Allia Bay; Astragalus; Machine learning; Human evolution

Resumo

A influência das alterações climáticas na evolução humana é muito debatida. Contudo, fósseis de hominídeos são escassos e diferentes direções de pesquisa, como métodos livres de táxon, tentam confrontar essa realidade. Este projeto visa aplicar tais métodos, utilizando várias técnicas de *machine learning* (ML), para reconstruir os aspetos paleoambientais de sítios africanos do Plioceno inicial associados ao *Australopithecus anamensis.*

O estudo proposto utilizou dados publicados de medições de astrágalos de bovídeos, retiradas de fósseis de Kanapoi (n = 31) e de Allia Bay (n = 15), e de indivíduos existentes (n = 187). Foram realizadas duas análises, a ecomorfologia, onde os indivíduos da amostra existente foram classificados em diferentes habitats e categorias de humidade do solo, e a ecometria, onde a amostra existente enquanto comunidade foi usada para prever várias variáveis de precipitação e temperatura e classificar as categorias de cobertura do solo. Finalmente, utilizando estes dados, vários algoritmos de ML foram treinados para cada análise, para inferir o possível ambiente da amostra fóssil.

Os resultados obtidos revelam melhores taxas de classificação e previsão ao aplicar diferentes algoritmos às análises, em comparação com a análise discriminante linear (LDA), método comumente utilizado nestas análises. Além disso, a análise ecomorfologica mostra habitats principalmente intermediários e uma mistura das categorias seca e húmida, para ambos os locais. No mesmo sentido, a análise ecométrica mostra locais com temperaturas tropicais (23,8 e 25,7ºC), húmidos (> 850 mm), mas muito variáveis e, consequentemente com vegetação altamente heterogênea, combinando florestas perenes de folha larga e bosques herbáceos, mas também sistemas mais abertos como pastagens. Estes dois sítios enquadram-se num contexto muito complexo e variável que pode contribuir para mudanças repentinas no ambiente. Isto parece indicar que *A. anamensis* prosperou nesses tipos de ambientes.

Em última análise, os métodos ecométricos parecem mais promissores, uma vez que os seus resultados permitem uma atribuição mais clara a um potencial clima. Assim, seria interessante estendê-los a outros locais importantes associados à nossa história evolutiva.

Palavras-chave

Paleoambientes; Kanapoi; Allia Bay; Astrágalo; *Machine learning*; Evolução humana

Acknowledgements

This project is the conclusion of an excellent academic journey, with exceptional teachers, who throughout these five years have allowed me to acquire profound knowledge on the important field of Human Evolution.

First of all, I would like to thank my supervisors, Prof. Dr Cláudia Umbelino, from the University of Coimbra and Dr Thomas Püschel, Postdoctoral Researcher at the University of Reading and Research Affiliate at the University of Oxford, for all their support and help. I am very grateful to Dr Thomas for his mentorship. It was a real pleasure to cross my academic path with him, his enthusiasm and passion for the subject and his excellent sense of humour make him a professional that I inspire one day to be.

A special thanks to DPhil Candidate João Coelho for all his help and willingness since the beginning of this project. To the forensic anthropology laboratory, especially to PhD Student David Navega, for the access to the tools I needed for the initial project, although it ended up not being the same given the impositions of the pandemic.

I would like to express my deepest gratitude to my greatest inspiration, the woman who gave me life and made me who I am today, Carmo Oliveira. My mother has always put me first, filled our home with love, joy and knowledge and never stopped me from being whatever I wanted to be.

To my best friend, Beatriz Juncal and my boyfriend, Bruno Costa, who have always offered me support and encouragement throughout this course and dissertation, thank you. They are my pillars, without whom I could not support myself.

Finally, I would like to thank the remaining friends and family who have always cared for me.

1) Introduction

Since Charles Darwin (1871) placed humans in an evolutionary structure, a central purpose of paleoanthropology has been to understand how environmental factors may have shaped the evolution of our early ancestors over the past 7 million years (Laporte & Zihlman, 1983; Vrba et al., 1995; Potts, 1998a, 1998b, 2007, 2013; Bobe et al., 2002; Bonnefille et al., 2004; deMenocal, 2004, 2011; Wynn, 2004; Behrensmeyer, 2006; Vrba, 2007; Kingston, 2007; Maslin & Christensen, 2007; Elton, 2008; Maslin & Trauth, 2009; Levin, 2015; Marean et al., 2015; Campisano et al., 2017). This is reflected in the set of hypotheses that have been emerging, proposing that environmental changes have been responsible for some aspects of human evolution, such as morphological adaptation to bipedalism, increased cranial capacity, behavioural adaptability, cultural innovations, cross-continental immigration events, among many others (e.g., Wheeler, 1991; Potts, 1998a; Bobe et al., 2002; Plummer, 2004; Plummer et al., 2009; White et al., 2009a; Braun et al., 2010; van der Made, 2011; Antón et al., 2014; Domínguez-Rodrigo, 2014; Böhme et al., 2019).

However, establishing the cause of the links between changes in the environment and the evolution of hominins is not a simple process. Data that directly relate organisms to their surroundings are available only for the present, where conditions are known and therefore possible to infer. Even though in recent decades there has been a better understanding of paleoclimate and palaeoenvironmental changes, capturing this knowledge and understanding which selective pressures shaped the evolution of our ancestors remains an ongoing challenge (Behrensmeyer, 2006; Kingston, 2007; Potts, 2007; Marean et al., 2015). For example, to address many paleoanthropological questions, reliable estimates of when hominin taxa originated and became extinct are needed (Du et al., 2020), although recent research has contributed towards this effort (e.g., Püschel et al., 2021). Nevertheless, hypotheses on the subject are difficult to test, and as a result, the impact of climate and environment on human evolution has been and will continue to be debated (Vrba, 2007; Elton, 2008).

It is now widely acknowledged that at the end of the Miocene a change in climate occurred and, globally, there was a cooling of the temperature extending across the Plio-Pleistocene (Zachos et al., 2001). Traditionally, the preferred interpretations propose that human morphological and behavioural adaptations had been influenced by the environmental pressure of an expanding dry savannah, in the light of temperature decreases (Vrba, 1995a;

Feakins et al., 2013; Uno et al., 2016; Will et al., 2021). The discovery of the first *Australopithecus*, a child's skull, in Taung (South Africa), as well as *Australopithecus bahrelghazali* from Koro Toro region (Chad) (Brunet et al., 1996), among other fossils indicative of a grazing environment, prompted different explanations that attempted to link these first hominins, open grasslands and the origins of bipedalism (White et al., 2009b).

Nevertheless, based on more recent environmental studies and discoveries, forests or more forested environments are now being considered (Pickford & Senut, 2001; Vignaud et al., 2002; Haile-Selassie et al., 2004; White et al., 2009a, 2009b; Passey et al., 2010; Cerling et al., 2011; Bamford et al., 2013; Roche et al., 2013; Senut, 2015). Thus, it is now accepted that their habitats were not restricted to a single type of environment, but rather a mixture of environments, woodland, and savannah, often described as a "mosaic" (Laporte & Zihlman, 1983; Kingston et al., 1994; Leakey et al., 1995; Plummer, 2004). This interpretation is associated with the heterogeneity of habitat (Reynolds et al., 2015; Barr & Biernat, 2020) and emphasizes the ability of our ancestors to adapt or respond to environmental changes (Carrión et al., 2019). This can be illustrated with early hominins, such as for example *Australopithecus sediba,* which shows post-cranial anatomy that seems indicative of both bipedal locomotion on the ground, as well as climbing trees (Berger et al., 2010).

However, these reconstructions are often inconsistent (Wood & Strait, 2004) with several different habitat scenarios being proposed. One reason for the lack of unanimity and multiplicity of hypotheses regarding the palaeoenvironment lies in the definitions, as shown in a recent review of the ‘Savannah hypothesis’ in human evolution studies (Domínguez-Rodrigo, 2014). According to this author, the term "savanna" is both described as a pure pasture and mosaic habitats with a substantial number of trees in a pasture environment, depending on the publication.

Part of this problem lies in the complexity of human evolutionary history, as hominin fossils are rare and fragmented, thus hominin palaeoecology is not providing more informative evolutionary insights (Faith et al., 2021). Besides, this oldest fossil hominin record has generally been found so far in areas covering only a portion of the African continent. The fossils are mainly derived from the Eastern branch of the East African Rift System (EARS) and caves in South Africa, where conditions for fossilization and subsequent exposure are favourable (Wood, 2002). Furthermore, in most of these cases, hominins are among the least abundant mammal fossils recovered from a locality (Bobe & Behrensmeyer, 2004; Levin, 2015), but are persistent (Villaseñor et al., 2020), hence contributing to the uncertainties surrounding this issue. This is partly explained by fossilization bias, as this process is highly complex (Kidwell & Behrensmeyer, 1988). For it to take place there has to be a substitution of organic matter for mineral material. However, not all types of deposition environments have the best conditions or even the same collection protocols, so it is quite unusual to find a fossil of a complete organism, particularly challenging for paleoanthropologists, who often deal with isolated teeth or specific skeletal parts such as mandibles (Senut, 2015).

This problem intensifies as we deal with older organic remains, thus several palaeoenvironmental questions related to the emergence of hominids during the Miocene and their early radiation through the Pliocene remain unanswered, such as for example the emergence of bipedalism in hominids (Böhme et al, 2019). Nevertheless, this period is crucial for understanding how environmental factors may have influenced the early evolution of our ancestors (Agustí, 2007; Elton, 2008).

The fossil record is, however, a very important and direct source of the history of life on Earth when the world existed in other circumstances, and new promising research directions are emerging to address more effectively the reality of an incomplete fossil record (Faith et al., 2021). Analysing this evidence to understand the relationship between ecosystem dynamics and the paleobiology of hominins is critical to test hypotheses that attempt to elucidate the influence of the environment on the morphological and behavioural adaptations of early hominins (Patterson et al., 2017). In addition to being one of the fundamental issues of palaeoanthropology, interest in the processes of environmental change has drawn greater attention from the general public.

Despite the difficulty, fieldwork conducted over decades has generated large collections of other fossils, including a solid record of mammals (Bobe & Leakey, 2009; Werdelin & Sanders, 2010; Bobe, 2011). This is primarily due to the fact that their teeth and skeletons withstand fossilization far better than those of most other organisms (Andrews, 2018). In fact, creating a complete picture of human evolutionary history requires an understanding of the animals and plants that have lived alongside the hominins, a knowledge of how they interacted with their environments (Polly et al., 2011). After all, the evolution of our ancestors did not take place in a closed ecological context and must therefore be incorporated into the same ecological ground as other animals that coexisted at the same time and space (Hardt et al., 2007).

Moreover, extinct mammals have long been a prime source of information about the ancient habitats in which they lived (Andrews, 2018; Kovarovic et al., 2018), and the analysis of the evolutionary context of vertebrates to study the evolution of hominids has emerged as a strong area of research (Vrba, 1995b; Behrensmeyer et al., 1997; Bobe & Behrensmeyer, 2004; Bobe et al., 2007a; Werdelin & Sanders, 2010). Particularly bovids are often used in these African paleo-ecological environmental reconstructions (Bobe, 2006; Bobe et al., 2007b; Barr, 2017). As a diverse family of ruminant ungulate mammals, bovids include antelopes, cattle, duikers, gazelles, goats and sheep, varying greatly in terms of mass, also varying in terms of habitat, occurring in a wide variety of environments, and feeding on a wide variety of plants (Castelló, 2016). Additionally, they are a dominant component of most African mammal fossil assemblages (Faith & Lyman, 2019), thus making them a particularly useful palaeoenvironmental indicator.

Analysis of this remarkable fossil record can provide essential insights into the interaction of environmental changes and mammalian evolution, including processes that are likely to affect hominin evolution (Bobe et al., 2007b). As a result, many studies have focused on these fossil mammals to understand changes in environments and climate, for example, using stable isotope records preserved in dental enamel (e.g., White et al., 2009b; Cerling et al., 2011; 2013; 2015; Sponheimer et al, 2013; Roche et al., 2013; Wynn et al., 2016; Blumenthal et al., 2017) and taxon diversity analysis (e.g., Behrensmeyer et al., 1997; Bobe & Behrensmeyer, 2004). However, many of these analyses are in part limited by the difficulty of taxonomic identification (Kingston& Harrison, 2007), especially difficult when dealing with fossil limb bones and other less-complete material (Barr, 2018).

Taking these problems into consideration, taxon-free approaches such as ecomorphology and ecometrics have been developed (Bock, 1989; Damuth, 1992; Ricklefs & Miles, 1994; Eronen et al., 2010a; Polly et al., 2011; Žliobaitė et al., 2016; Vermillion et al., 2018). They have served as important sources of evidence for the environmental context of human evolution (Barr, 2015, 2018).

1.1 Taxon-free methods

Taxon-free methods comprise a family of techniques that links the present and past through the functional traits of species, disregarding their shared evolutionary origin (Damuth, 1992; Polly et al., 2011; Andrews & Hixson, 2014**)**. This premise is supported by the convergent evolution of ecomorphological traits between distantly related species groups (Andrews & Hixson, 2014) and phylogenetically unrelated taxa (e.g., Barr, 2014). This is particularly a critical requirement to compare changes in the modern world with those of the distant past, as species are constantly becoming extinct and new species originate.

These methods look for morphological traits that have a functional relationship with the environment (i.e., functional traits) and ultimately this relationship is applied to fossils to infer ecological traits of extinct taxa (Eronen et al., 2010a; Polly et al., 2011; Žliobaitė et al., 2016; Barr, 2018; Vermillion et al., 2018; Faith & Lyman, 2019). Ecomorphology deals with the study of this complex trait-environment dynamics in individuals (Barr, 2017, 2018) and ecometrics, rather than focusing on individual organisms, deals with the functional composition of communities (Žliobaitė et al., 2016; Vermillion et al., 2018). Therefore, they are not mutually exclusive (Faith & Lyman, 2019).

That said, it is possible to use mammalian fauna, or any other taxon, that coexisted with early hominins to infer their habitat, as they responded to the same selective pressure as early hominins, although not in the same way. As organisms interact with their environment through their functional traits (Lawing et al., 2012; Barr & Biernat, 2020), a certain combination of biotic and abiotic conditions will consequently favour traits that maximise an organism's performance by natural selection (Vermillion et al., 2018). These functional traits will constrain a species' ability to remain and inhabit a particular habitat over time, regardless of its taxonomic identification (Barr, 2018).

However, the fact that many features are often phylogenetically restricted (Barr, 2018), for example, teeth are limited to vertebrates, it means that taxon-free approaches are not completely taxon-free. Nevertheless, comparisons between different communities without common species are still possible (Vermillion et al., 2018). This requires sufficient general correspondence between the function-environmental relationship to apply to any taxon in which the characteristic is found (Polly et al., 2011), meaning that species in different clades have adapted to the same habitat several times, consequently, extrapolating out of the sample can be done with relative confidence (Barr, 2018).

History of Ecomorphology

The first examples of taxon-free approaches to reconstruct the ecology of individual taxa are as old as palaeontology itself, with different naturalists, observing that an animal's form was related to the ecological context of its function (Faith & Lyman, 2019). However, the term ecological morphology, today known as ecomorphology, was only coined in 1948, by the Dutch zoologist Cornelis van der Klaauw.

There is now a substantial body of literature exploring how morphological traits can be used to understand aspects of diet (Janis, 1988; Van Valkenburgh, 1988; Figueirido et al., 2009; Meloro et al., 2015; Forrest et al., 2018) and locomotor preferences based on habitual substrate use (Van Valkenburgh, 1987; Elton, 2002; Polly, 2010; Meloro et al., 2013; Elton et al., 2016; Püschel et al., 2018, 2020), or based on predator avoidance strategies (Kappelman, 1988, 1991; Plummer & Bishop, 1994; Kappelman et al., 1997; Scott et al., 1999; DeGusta & Vrba, 2003, 2005; Kovarovic & Andrews, 2007; Plummer et al., 2008, 2015; Barr, 2014, 2015; Curran, 2012, 2015).

Whilst it is possible to develop environmental inferences from a free taxon-characterization of single species taxa, when functional characteristics are analysed at the community level, variation at the individual and species level tends to be mitigated, therefore tracking environmental gradients more closely than characteristics at the individual or species level (Polly, 2010).

History of Ecometrics

Ecometrics is a computational methodology focusing on functional characteristics of fossil communities (Eronen et al., 2010b; Polly et al., 2011; Vermillion et al., 2018). Although the term is relatively recent, ecometric analysis has a long history in palaeoenvironmental reconstruction and is used extensively today.

Its roots date back to early palaeobotany in the 20th century, with Bailey and Sinnott (1915) conducting the first study on the variations in the shape of the community dicotyledonous plant leaf relative to climate. Since then, different characteristics have been explored for ecometric analysis and strong ecometrics relationships have been documented for a range of characteristics and taxonomic groups, including leaf size and shape (Wolfe, 1995; Traiser et al., 2005), dental morphology of large herbivores (Fortelius et al., 2002, 2016; Eronen et al., 2010b, 2010c; Liu et al., 2012; Žliobaitė et al., 2016), limbs and locomotion (Polly, 2010; Polly & Head, 2015; Barr, 2017; Short & Lawing, 2021), skeletal traits (Lawing et al., 2012), as well as body mass (Meloro & Kovarovic, 2013). Some examples are provided below (Figure 1).

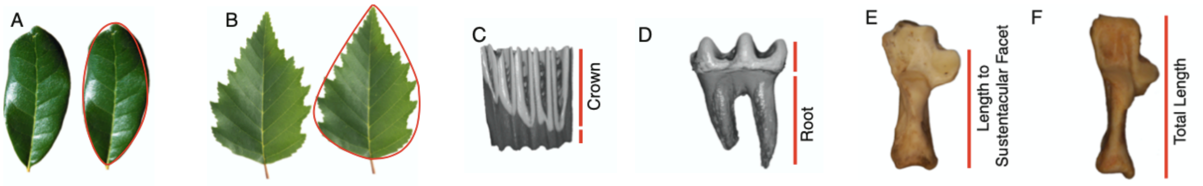

Figure 1 - Examples of the morphology of ecometric traits. A, Example of an entire leaf margin, characteristic of hotter environments. B, Example of a non-entire leaf margin, characteristic of cooler environments. C, Example of hypsodonty (high-crowned teeth) in a rodent, characteristic of arid environments with more gritty vegetation. D, Example of brachydonty (opposite to hypsodonty) in a rodent, found in more wet environments with mixed vegetation. E, Example of raccoon calcaneum. The gear ratio (ratio of the length of the sustentacular facet to the total length of the calcaneum) is characteristic for animals living in high vegetation cover (e.g., dense woodlands). F, Example of a panther calcaneum. The gear ratio in the panther is characteristic of animals that are adapted to run through open habitats (Adapted from Vermillion et al. (2018)).

This type of study has been growing in recent years and appears as an opportunity to unite different disciplines for the study of the biology of climate change (Polly et al., 2011). In analyses of faunal communities, hypsodonty (the ratio of tooth crown height to tooth root height) (Janis & Fortelius, 1988) is the most common feature and the first to be used as a proxy for vegetation structure (Fortelius et al., 2002). Crown height relates to precipitation because higher-crowned teeth (i.e., hypsodont teeth) is common in pastures and rare in temperate forests, since higher-crowned teeth enable animals to tolerate high levels of environmental grit associated with arid environments (Fortelius et al., 2002).

This selective pressure at individual tooth height reaches the level of the mammalian community and, therefore, average hypsodonty of mammals is strongly related to annual precipitation and has been demonstrated at a range of spatial scales, ranging from regional (Žliobaitė et al., 2016), continental and sub-Continental (Fortelius et al., 2016; Žliobaitė et al., 2018; Faith et al., 2019), to global (Eronen et al., 2010c; Liu et al., 2012).

Since measuring functional traits is often possible for both existing and fossil organisms, this approach provide an important tool for understanding the trait-environmental dynamics. In this regard, it is important to select traits with reference to a specific functional hypothesis that links them to an ecological variable of interest. Such traits can be typical linear measurements (e.g., Kappelman, 1988; Plummer & Bishop, 1994; Kappelman et al., 1997; DeGusta & Vrba, 2003; Plummer et al., 2008, 2015) or more complex variables such as those derived from geometric morphometrics (e.g., Figueirido et al., 2009; Cooke, 2011; Curran, 2012, 2018; Cano et al., 2013; Forrest et al., 2018; Püschel et al., 2018, 2020; Etienne et al., 2020).

The astragalus (or talus) (Figure 2) is typically among the most abundant postcranial elements in fossil collections (Hussain et al., 1983) due to its high density and corresponding resistance to hydraulic transport (Behrensmeyer, 1975). Additionally, the astragalus is also the main mechanical connection between the leg and foot and is responsible for transmitting body weight, as well as providing stability and mobility throughout locomotor behaviours (Boyer et al., 2015). This combination of its high occurrence and good preservation in the fossil record and its functional role in the ankle joint make it a good element to study possible morphological variations that may be linked to some ecological variable of interest (Püschel et al., 2018).

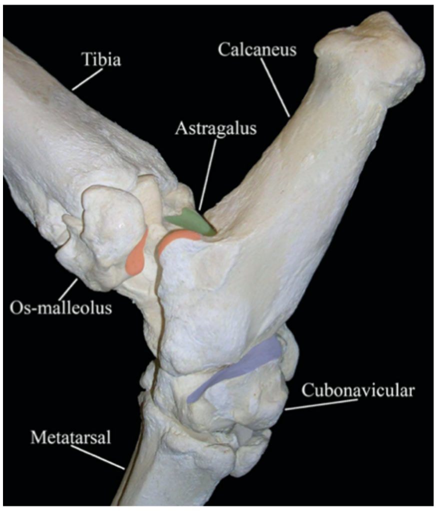

Figure 2 - Lateral view of left-side hock joint in articulation, with skeletal elements labeled and joint surfaces highlighted: astragalus- calcaneus joint (vertical green and black stripes); os malleolus-calcaneus joint (pink and black stippling); cubonavicular-calcaneus joint (blue and black horizontal stripes). Orientation abbreviations: A, anterior; I, inferior; P, posterior; S, superior. (Adapted from Curran,2018)

<u>*Astragalus: As a functional trait*</u>

As a taxonomic family, all bovids exhibit skeletal features linked to cursoriality in their common ancestry (Schaeffer, 1947). However, there are differences in the ankle joint anatomy of these bovids, since each habitat has different levels of vegetative complexity, which in turn require different forms of locomotion and predator avoidance behaviours (e.g., DeGusta &

Vrba, 2003; Plummer et al., 2008; Barr, 2014). For example, open habitat species must be fast and agile runners (Jarman, 1974), whereas closed habitat species presumably rely more on camouflage as they need to navigate a substrate such as dense forests that includes undergrowth and other obstacles (Kappelman, 1988; Kappelman et al., 1997; Plummer et al., 2008).

To understand the role of the astragalus, one must understand its function in the movement of the hock joint, as it is clearly explained in Barr's (2014) article. When the animal gives an impulse to move, during the early stages of plantarflexion (Figure 3A) the astragalus will rotate and the cavity between the calcaneus and cubonavicular must accommodate the rotating astragalus, which is ellipsoid and elongated in the anteroposterior dimension. Thus, astragalar rotation causes the cavity to expand, displacing the calcaneus posteriorly and most of the functional length of the astragalus is added to the effective length of the calcaneus. This movement is complete when the calcaneus, astragalus and cubonavicular reach a closed position (Figure 3G).

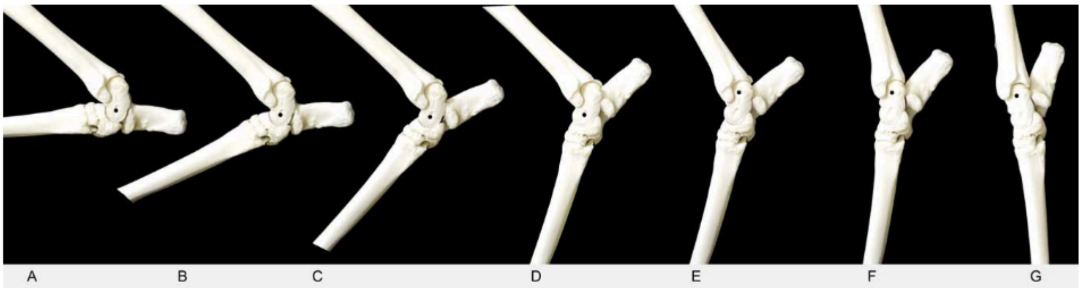

Figure 3 Photographs of Ammotragus lervia hock joints in medial view in various stages of the stride cycle from extreme dorsiflexion (A) to extreme plantarflexion (G). The operative center of rotation of the joint is indicated by a black dot on the astragalus. (Adapted from Barr (2014)).

Given this, Barr (2014) concludes that forest bovids are characterised by longer and narrower astragalus. This morphology has a dynamic mechanical advantage for the plantarflexion of the hock joint, as more proximo-distally elongated astragali act to displace the calcaneus even further posteriorly during plantarflexion and increase muscle strength during this movement, producing a powerful plantarflexion (Barr, 2014), following the principles of mechanical lever systems[1] (Figure 4).

---

[1] There are three types of lever systems, Barr (2014) compares the hock joint to the first class lever system, which he describes as consisting of an lever (force) arm (calcaneus), a load (resistance) arm (length of the remaining tarsals, metatarsals and phalanges) and a fulcrum or center of rotation (astragalus), which is between the force and resistance. The relative proportions of the lever and load arms will affect the relative speed and power of this first-class lever system, i.e. increasing the relative length of the lever arm will result in a more powerful but slower lever system and the opposite creates a faster but less powerful system (Barr, 2014).

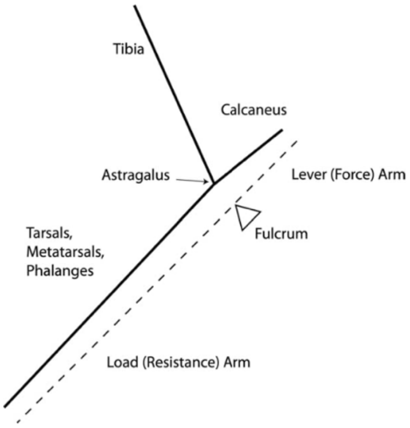

Figure 4 Schematic diagram of the bovid hock joint as a simple first-class lever system during plantarflexion. Note: this diagram is simplified, and does not consider the multiple centers of rotation that are present in the bovid astragalus. (Adapted from Barr (2014)).

On the other hand, bovids occupying pasture are characterised by shorter and wider astragals, which causes less posterior displacement of the calcaneus during astragalar rotation, reducing the lever arm force for the plantarflexion of the hock joint, but increases the range of motion and foot speed during locomotion (Barr, 2014), according to Muller's (1996) four-bar linkage model[2].

Furthermore, open-field cursors are predicted to increase the surface areas of the astragalar joint relative to body size (Barr, 2014) because the dissipation of loads over a larger surface area reduces the risk of joint damage (Jungers, 1991).

Hence, astragalus is strongly related to vegetation structure and this relationship has been well studied (Plummer & Bishop, 1994; DeGusta & Vrba, 2003; Plummer et al., 2008, 2015; Barr, 2014, 2015, 2017) and can be useful for reconstructing the palaeobiology and palaeoenvironment of extinct animals (Barr, 2017).

As vegetation productivity in most terrestrial environments depends on humidity and temperature, and other factors such as fire, precipitation seasonality and soil characteristics (Bond, 2008; Good & Caylor, 2011), in the absence of anthropogenic actions (Mayaux et al., 2004), it is likely that one can estimate these climate variables directly from these locomotor characteristics, due to their relationship with the structure of the vegetation. This had been previously shown, with Barr's (2017) study applied to present-day sub-Saharan African bovids,

---

[2] According to this model, a relatively shorter astragalus should result in a greater potential range of motion, leading to longer potential stride lengths, and therefore increases maximum running speed, while a relatively longer astragalus should result in shorter and therefore slower potential stride lengths.

revealing that astragalus explain large proportions of variation in annual precipitation and land cover ($R^2 > 0.6$).

As opposed to ecomorphological studies (e.g., Kappelman, 1991; Plummer & Bishop, 1994; Kappelman et al., 1997; Kovarovic & Andrews, 2007; Plummer et al., 2008, 2015; Kovarovic et al., 2021; Dumouchel et al., 2021), these ecometric methods have not been widely applied to the African continent's fossil mammalian registry (Fortelius et al., 2016; Žliobaitė et al., 2018), nor have they been extensively explored for the astragalus bone, although they have proven extremely useful for understanding biological responses to climate change (Eronen et al. 2010a).

Besides, the majority of ecomorphological studies have been carried out using linear classification methods (Curran, 2018), such as the popular linear discriminant analysis (LDA), a multivariate technique designed to classify individuals (i.e., specimens) into predefined categories based on an observed set of measured variables (e.g., Kappelman, 1988, 1991; Plummer & Bishop, 1994; Kappelman et al, 1997; DeGusta & Vrba, 2003, 2005; Bishop et al., 2006; Kovarovic & Andrews, 2007; Colangelo et al., 2010; Cooke, 2011; Fraser & Theodor, 2011; Curran, 2012, 2015; Cano et al., 2013; Plummer et al., 2015; Forrest et al. 2018; Kovarovic et al., 2021), despite the known limitations of these approaches (for further details about these limitations see for e.g., Feldesman, 2002; Mitteroecker & Bookstein, 2011).

Therefore, the proposed research will generate ecomorphological and ecometric models based on a series of machine learning algorithms trained using morphometric data from extant astragalus applied to a fossil sample to reconstruct the palaeoenvironments of the African earliest Pliocene. Although the application of machine learning algorithms to address specimen identification or group characterization problems has a vast literature in other biological fields (Tarca et al, 2007), only more recently have several machine learning methods been applied using morphometric data (e.g., MacLeod, 2007, 2017; van Bocxlaer & Schultheiß, 2010; van den Brink & Bokma, 2011; Navega et al., 2015; Sonnenschein, 2015; Li et al., 2016; Hanot, 2017). Besides, most of them did not compare different algorithms applied to the same problem (Püschel et al., 2018). Nevertheless, incorporating ML-derived predictive modelling techniques into the functional morphology toolkit may prove to be a useful addition, offering more flexibility and predictive power (Püschel et al., 2018).

The models will be based in the Kanapoi and Allia Bay sites, located in the Lake Turkana region of Kenya's Rift Valley.

1.2 Geological context

East African Rift System (EARS)

Most of the key events of early human evolution (before ~1 Myr) occurred in Africa (MacLatchy et al., 2010). Based on patterns of the temporal and spatial distribution of hominins resulting from fossil recoveries, the East African Rift System (EARS) and South Africa have often been considered not only as fossil-rich sites but also as 'birthplaces' therefore being taken as the principal theatre of hominin evolution (Maslin et al., 2014). It comprises the earliest occurrences of genera and most of the earliest for fossil hominin species, including *Ardipithecus*, *Orrorin*, *Australopithecus*, *Kenyanthropus*, *Paranthropus* and *Homo* (Carrión et al., 2019). As a result, focusing on these areas to reconstruct the climate context provides an understanding of the early hominin environment background.

The East African Rift System (EARS) (Figure 5) is one of the most extensive, running north-south for approximately 4,500 km from Syria through East Africa to Mozambique (Maslin & Trauth, 2009). It is densely fractured and consists of four main fissure zones: Afar, Rift Main Ethiopia (MER), Kenya-Tanzania and Western Clefts (WoldeGabriel et al., 2016).

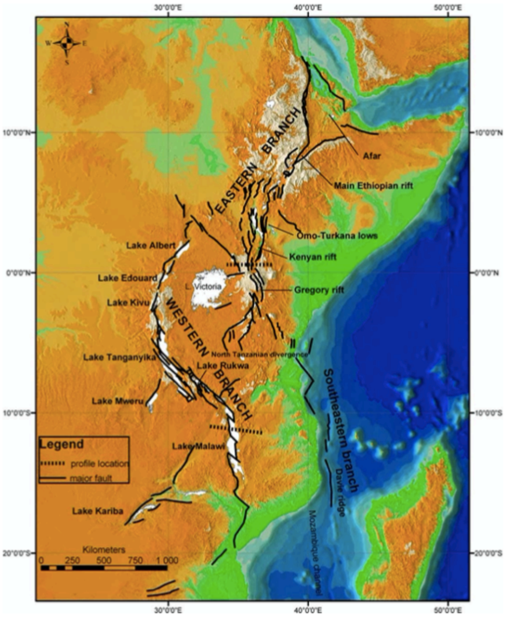

Figure 5 - Hypsographic DEM (Digital Elevation Model) of the East African rift system. Black lines: main faults; white surfaces: lakes; grey levels from dark (low elevations) to light (high elevations) (adapted from Chorowicz (2005)).

Rifting is one of the basic principles of plate tectonics, as it explains the dismantlement of continents and the development of oceans (Scoon, 2018). The Kenya Rift extends from Lake Turkana, in northern Kenya, to the northern tip of Lake Natron, on Kenya's border with Tanzania (WoldeGabriel et al., 2016).

Turkana Basin

Lake Turkana is located in northern Kenya, but the Turkana Basin is part of the wider geological context that includes the Omo valley in Ethiopia (Bobe, 2011). This region is especially well known for the fossil presence of *Australopithecus anamensis* (Feibel et al., 1991; Haile-Selassie et al., 2019) at Kanapoi and Allia Bay (Figure 6), when it was first named and described (Leakey et al., 1995, 1998; Ward et al., 2001). Currently is considered to be the earliest species of the genus (Bobe et al., 2020).

Figure 1. Geographical map showing the location of Allia Bay and Kanapoi within the Omo-Turkana Basin and Africa.

Figure 6 - Schematic map of the Omo–Turkana Basin, including Lake Turkana and the Lower Omo Valley. Inset shows the Turkana Basin in the context of eastern Africa (adapted from Dumouchel et al., (2021)).

The largest sample of *A. anamensis* (4.2 Ma - 3.8 Ma) (Bobe et al., 2020) derives from Kanapoi, currently with 74 individuals attributed to this species, including males and females, adults and juveniles, and cranial and postcranial elements (Ward et al., 2001; Ward et al., 2020).

Kanapoi

Kanapoi is a paleontological site in Kenya's Rift Valley, southwest of Lake Turkana. Geological and paleontological research in this region began in 1964, conducted by Bryan Patterson of Harvard University (Patterson & Howells, 1967). During this expedition, Patterson and his colleagues discovered a distal part of a hominin humerus (KNM-KP 271) (Patterson & Howells, 1967). Later, from 1994 to 1997, a team of Kenya's National Museums (NMK) led by Meave Leakey returned to the fieldwork in Kanapoi, documenting important fossil hominins and naming a new species, *Australopithecus anamensis* (Leakey et al., 1995, 1998; Ward et al., 2001). This work in the 1990s resulted in the recovery of 650 vertebrate fossils (Harris & Leakey, 2003). In the following years, more fossils were recovered and, consequently, Kanapoi has the largest sample of vertebrate fossils in Africa contemporaneous to *Australopithecus anamensis* (Bobe et al., 2020).

Kanapoi formation deposits consist of a lower fluvial sequence, interrupted by a lacustrine episode, followed by an upper fluvial sequence limited by the Kalokwanya Basalt (Leakey et al., 1995, 1998; Feibel, 2003). Most sediments of Kanapoi Formation have accumulated between 4.3 and 3.9 Ma, but some sedimentation continued until 3.4 Ma, the age of the Kalokwanya Basalt that closes the sequence (Feibel, 2003). The most fossiliferous range of Kanapoi is limited in time (~ 4.2 - 4.1 Ma) (Leakey et al., 1995, 1998; Feibel, 2003) and its sediments appear in a limited spatial extent (~ 25 km2) (Feibel, 2003)

Fauna-based reconstructed environments indicated open and closed environments (Harris et al., 2003; Sanders, 2019) as did analyses of stable paleosol isotopes, that produced evidence for wooded and grassy environments (Wynn, 2000; Quinn & Lepre, 2020). The latter study's result supports environmental heterogeneity rather than temporal averaging where mixed habitat faunas exist due to the average time of distinct habitats within the relatively short time window (e.g., Behrensmeyer et al., 2007).

Allia Bay

Allia Bay is a region on the east side of Lake Turkana in Kenya. The first exploration of the Koobi Fora Formation took place in 1968, where it was recognised that this locality contained fossils (Harris, 1983). These fossils were collected from the surface over the following decades. Coffing and collaborators (1994) described these initial surface finds and attributed them to *Australopithecus* cf. *A. afarensis*. It was not until 1995 that excavations took place at the Allia Bay 261-1 site, carried out for three seasons until 1997 (Ward et al., 1999). These excavations resulted in the recovery of thirty-one fossil remains attributed to *A. anamensis* (Leakey et al., 1995, 1998), consisting mostly of isolated teeth (Behrensmeyer & Reed, 2013).

During the last field season over 2,000 fauna remains were recovered during the sieving process. Most of these fossils were unidentifiable and fragmentary non-mammals and are described in Hagemann (2010). The bone bed itself consists mainly of bones of aquatic species, mainly fish and reptiles. Many of the fossils in this bed are rolled and heated, indicating transport from a distance, reworking of older sediments, or both (Ward et al., 1999). Unlike Kanapoi, Allia Bay is more time restricted (Behrensmeyer & Reed, 2013) and Allia Bay fossils are slightly younger than those from Kanapoi (Ward, 2001).

The site is a fluvial deposit, represented by a channel filled with medium to coarse sand, with occasional fine sands and lenticular lamite interbeds, which form the fossiliferous sediments reaching a local thickness of at least 14 metres (Coffing et al., 1994). The primary structures are poorly preserved, but there is a 20 cm bone bed at the site, where most hominid teeth have been found. This is a lag deposit in the channel, about 5 m below the base of the Moiti Tuff, 1 m thick, dated elsewhere in the basin to 3.97 ± 0.03 Ma (Coffing et al., 1994; McDougall & Brown, 2008).

Environment reconstructions for Allia Bay based on isotopes indicated seasonal and open (Macho et al., 2003) and mosaic habitat of closed woodland and grasslands with higher rainfall than the region receives today (Schoeninger et al., 2003). The fauna of Allia Bay suggests that the local habitat probably included a mixture of habitats, with more closed gallery forest environments, but also lowland grasslands and marginal dry woodland (Feibel et al., 1991; Coffing et al., 1994; Leakey et al., 1995; Schoeninger & Reeser, 1999). Geological evidence supports this palaeoenvironmental reconstruction and suggests that the area exhibited seasonal fluctuations in moisture (Brown & Feibel, 1991; Wynn, 2000).

These sites hold a special place in palaeoanthropology because of their unique record of hominin evolution in the Plio-Pleistocene (Fleagle & Leakey, 2011). Consequently, they have been extensively studied and contain a solid database of various animals, especially bovids.

Hence, developing ML models and applying to these specimens will be useful in assessing the impact of climate change on the emergence of the first hominins. These models can subsequently be applied to any location of interest with materials from that location.

1.3 Aims

1) To reconstruct some of the palaeoenvironmental aspects of the Early African Pliocene by performing ecomorphological and ecometric analyses using fossil bovid astragali, testing various machine learning methods to better predict different environmental variables that have not yet been tested with astragalar data.

2) To relate the palaeoenvironmental reconstructions obtained with the emergence of early hominids to provide ecological contexts for a key episode in our own evolutionary history.

2) Materials and Methods

2.1) Sample

The main data source for this study comes from Dumouchel et al. (2021). In their paper, they performed a multiproxy analysis combining astragalar ecomorphology, dental mesowear, hypsodonty index and carbon and oxygen isotopes of mammalian tooth enamel to reconstruct the palaeoenvironment of Allia Bay, while comparing it with the results obtained from Kanapoi. For the purposes of this study, it was intended to focus only on the data used in the ecomorphology of the astragalus.

Astragalar data

The astragalar data comprises 15 fossil individuals from Allia Bay and 31 from Kanapoi, both held in the National Museums of Kenya, Nairobi, Kenya (detailed in Appendix A, Table A1). The data from these specimens was collected from published papers (Geraads & Bobe, 2020; Geraads et al., 2013; Harris et al. 2003). Unfortunately, due to the inherent difficulty associated with the taxonomic identification of post-cranial materials, none of the specimens was identified to the tribe rank.

To conduct an ecomorphological analysis it is necessary to take measures of the functional characteristics of the species (Vermillion et al., 2018). Dumouchel and colleagues (2021) achieved this by collecting five linear measurements for the astragalus of each specimen, following the work of Barr (2015) (Table 1). These linear measurements capture aspects of astragalus shape and size (Figure 7).

Table 1 Measurements used in this study. All measurements are defined in further detail in Barr (2014).

Measurement	Description
B	Functional length of the astragalus
DistRad	Radius of circle fit to the margin of the distal trochlea
MML	Maximum medial length
ProxRad	Radius of circle fit to the margin of the proximal trochlea
WAF	Width at level of flange on lateral surface, but excluding flange

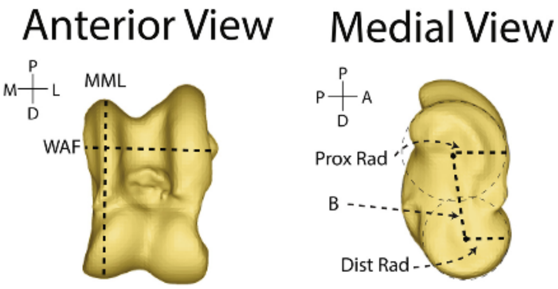

Figure 7 Anterior and medial views of a bovid astragalus, illustrating the five linear measurements used (adapted from Barr (2015)).

Furthermore, analysing this data requires comparison with extant data. Hence, a comparative sample was generated by using the data from Barr (2014). In this paper a diverse sample of 187 individuals from 53 extant bovids species (detailed in Appendix A, Table A1), of the 55 species of bovid that occur in sub-Saharan Africa (Barr, 2017), was used to link astragalar morphology with habitat categories.

These individuals are curated in the osteological collections of the American Museum of Natural History (AMNH), New York, NY, and the National Museum of Natural History (NMNH), Washington, DC. Even though Barr (2014) took 14 measurements for each one of the analysed individuals, it was only used the five measurements that were also available for the fossil specimens. All measurements were collected following the same method.

2.2 Methods

Data analysis was carried out using different machine learning algorithms. This discipline combines artificial intelligence, statistics, and data science to deal with topics, such as algorithm development for classification, prediction and pattern recognition based on models derived from existing data (Jordan & Mitchell, 2015).

The main decisions in building predictive modelling algorithms includes data preparation, data splitting, selection of input variables, as well as evaluation and model selection procedures. Data preparation is a critical step when building a successful machine learning model (Zhang, 2008). Therefore, the first thing that is important to ensure is the quality of the data.

After compiling the data from Dumouchel et al. (2021) (i.e. the fossil data), with the data from Barr (2014) (i.e., the extant data), an initial exploration of the data was carried out.

The first analyses were done in PAST 4.04 (PAleontological STatistics) software (Hammer et al., 2001), a user-friendly program, including a wide range of both general and special methods used within the field, but naturally without all the features found in more specialized programs.

A Principal component analysis (PCA) was performed to summarize the variance of this sample. PCA is a descriptive and explorative method that does not make any statistical assumptions (Hammer & Harper, 2006). It is a way of projecting points from the original, high-dimensional variable space onto a two-dimensional plane, with a minimal loss of variance (Hammer & Harper, 2006). It helped not only in understanding how the variables correlated with each other and the way they contributed to the analysis but also in detecting the existence of potential outliers. An outlier is generally considered as an observation that is significantly distant from the other observations considered (Benatti, 2019). Since there is no access to the original database, it was not possible to rectify the measurements and therefore, there was the need to eliminate these specimens from the analysed dataset.

2.2.1) Ecomorphological analysis

In an ecomorphological analysis, the first essential task is to associate environmental conditions of interest with the comparative species (Barr, 2018). This means that the goal is to achieve a correspondence between these ecological/environmental conditions and these specimens' measurements which characterize their morphology. Later this correspondence will serve to classify the fossils, based on their measurements.

Since the extant species had already been used for an ecomorphological analysis in Barr (2014), they had a designated habitat category. Hence, it was used the same categorization scheme, but some of the specimens needed to be classified as there were no categories for them in the original dataset (*Ammotragus lervia, Rupicapra rupicapra,* and *Bubalus bubalis*). Categories were assigned to these specimens by carrying out a literature review (detailed in Appendix B) while following the same classification as Barr (2014):

`Forest´ (F) bovids occupy environments where the tree canopy is closed, including rain forest and temperate forest. `Open´ (O) habitat bovids include those occupying grasslands, as well as more arid environments. `Heavy Cover (HC)´ habitats include bushland and woodland, as well as swampy habitats. `Light Cover (LC)´ includes bovids from light bushland, as well as bovids that specialize in tall-grass habitats.

In addition, to further capture more information, a new category, `Wet and Dry´, was added, considering that soil moisture is determined by temperature, precipitation, and atmospheric partial pressure of CO2 (PCO2) (Iio et al., 2013), i.e., factors that can certainly influence vegetation type. Finally, resorting again to the literature it was possible to categorize the soil type where the species live as `Wet or Dry´ (detailed in Appendix B).

In a certain way, these categories are already implicit in the original habitat categorization. However, the focus was not on the vegetation type, but more on soil moisture. Thus, the `Dry´ category includes habitats with a wide variety of vegetational cover, from open grasslands to temperate forests, whereas the species in the `Wet´ substrate category are also found in a variety of habitats, from swamps to tropical forests.

The remaining analyses, used to evaluate the effect of the environment on astralagar morphological variation, were conducted in RStudio v1.3.1073 using the *Caret* package v.6.0-88. (Kuhn, 2021), which consist of a set of simple and easy to write functions that help to streamline the generation of complex predictive models. With this package, it was possible to prepare and run different classification models (i.e., supervised algorithms). Classification is the task of predicting a discrete class label, it may predict a continuous value, but it has to be in the form of a probability for a class label (Kuhn & Johnson, 2013). These are supervised algorithms because labelled datasets are provided to the machine that will be then used to train the models. In this case the labels are the categories.

Prior to the analysis, all variables were size-standardized, so the data was scaled by geometric mean (defined as the average of a set of products) using Microsoft Excel 16.47 and transformed in PAST using Isometric Burnaby's method (Burnaby, 1966) (detailed in Appendix A, Table A1). This last size correction method projects all the set of measured distances onto a space orthogonal before analysis resulting in retention of consistent shape (same proportions) over a range of sizes (Hammer & Harper, 2006). This process seeks to assess the effect of species size on the results, given that it is natural for size to become a confounding factor when dealing with morphometric data. A confounding factor in a study is a variable that is related to one or more variables defined in that study and can mask a real association or falsely demonstrate an apparent association between study variables when there is no association between them (Skelly et al., 2012).

Therefore, a principal component analysis (PCA) was again performed, this time using the raw data, but also the data scaled by geometric mean, and data transformed using Burnaby's.

To represent a wide range of different classification models, models arising from different families will be tested, among them: (i) linear discriminant analysis (LDA); (ii) classification and regression tree (CART); (iii) k-nearest neighbours (KNN); (iv) naive Bayes (NB); (v) support vector machine (SVM) and (vi) random forest (RF). Below a brief description is provided for each one of these methods.

(i) linear discriminant analysis (LDA);

The linear discriminant analysis developed by Fisher (1936) and Welch (1939) stands out from the various methods available within the discriminant analysis family. In different but equivalent approaches, Fisher (1936) proposed that the idea is to find the linear combination of variables so that the variance between groups is maximized relative to the variance within the same group. This means that the goal is to find the combination of variables that provides maximum separation between data centres, while minimizing the variance within each data group (Fisher, 1936; Welch, 1939).

(ii) classification and regression tree (CART);

CART models are achieved by recursively partitioning the data space and fitting a simple prediction model within each partition (Loh, 2011). As result of this recursive process, the partitioning can be graphically displayed as a decision tree (Breiman, 1984). Classification trees are intended for dependent variables that take a finite number of unordered values (i.e., categorical categories), with prediction error measured as misclassification cost. The aim of CART is to partition the data into smaller, more homogeneous groups (Breiman, 1984).

(iii) k-nearest neighbours (KNN);

KNN assumes that similar things exist very close to each other and therefore uses feature similarity to classify observations (Venables & Ripley, 2013). When KNN is applied for classification, it predicts the class of an observation using the K-closest samples from the training set, where K refers to the number of neighbours that were used to determine class membership (Peterson, 2009). The value of K can be set arbitrarily by the user or adjusted by applying different approaches (e.g., using a grid search). The "proximity" between observations is established by computing a distance metric (e.g., Euclidean, Minkowski, among many others), while the choice of this metric depends on the characteristics of the variables (Venables & Ripley, 2013). Regardless of the calculated distance metric, it is essential to remember that the original scales of the variables affect the resulting distance calculations (Püschel et al., 2018). Therefore, to allow each predictor to contribute equally to the distance computation, it is recommended to centre and scale all variables before performing KNN (Kuhn & Johnson, 2013).

(iv) naive Bayes (NB);

NB is a relatively simple technique for generating classification models that allocate observations according to class/category membership to problem instances, represented as vectors of feature values, where class membership is drawn from some finite set (Kuhn & Johnson, 2013). NB is based on Bayes' rule, which basically computes the probability that the outcome belongs to a certain class membership, given a set of variables (Webb et al., 2005).

The core of the model is the estimation of the conditional and unconditional probabilities associated with the set of variables (Püschel et al., 2018). Despite their naive design and simplified assumptions, NB classifiers have performed surprisingly well in many complex real-world situations (Caruana, 2006).

(v) support vector machine (SVM)

Support vector machines are a family of statistical models that have evolved considerably (Kuhn & Johnson, 2013) and are considered one of the most successfully applied techniques for pattern recognition (Decoste & Schölkopf, 2002). In brief, SVM tries to define what is the best decision boundary to separate classes given the existence of many possible decision boundaries that can separate all the training samples into classes correctly (Marcé-Nogué et al., 2017). Among all these possible decision boundaries, SVMs find the one that achieves maximum margin between the categories. The margin is defined as the distance between a planar decision surface that separates the classes and the closest training samples to the decision surface. SVMs are part of a wider group of techniques known as kernel methods that owe their name to the use of kernel functions (Hastie et al., 2017). These functions enable these algorithms to work in a high-dimensional space without estimating the coordinates of the data in that space, but rather by just computing the inner products between points in a suitable feature space, which reduces computational cost even in very high-dimensional spaces (Ben--Hur et al., 2008).

(vi) random forest (RF).

RF or random decision forests are an ensemble learning method for classification and regression that work by generating several decision trees during training, to then output the class that is the mode of the classes (i.e., classification) or mean prediction (i.e., regression) of the individual trees (Kuhn & Johnson, 2013). In simpler words, RF is a machine-learning method that generates multiple decision trees and merges them together to get a more accurate, robust, and stable prediction (Püschel et al., 2018). RF are a combination of tree predictors/variables such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Lin & Jeon, 2002). Each one of the trees in the forest 'votes' when classifying of a new observation/sample, and the proportion of votes in each category across the ensemble is the predicted probability vector. Whereas the kind of tree changes in the algorithm, the tuning parameter of number of randomly selected variables to choose from at each split is the same (Ho, 1995).

The performance of the classification predictions was evaluated using overall accuracy (i.e., error rate), as well as Cohen's Kappa coefficient (Püschel et al., 2020). A combination of different metrics provides a complete picture of error distribution (Chai & Draxler, 2014). The overall accuracy is a simple metric that reflects the agreement between the observed and predicted classes and has the most straightforward interpretation (Kuhn & Johnson, 2013). Cohen's Kappa statistic can range between −1 and 1, where a value of 0 means that there is no concordance between the observed and predicted classes, whilst a value of 1 would indicate perfect agreement of the model prediction and the observed classes (Kuhn & Johnson, 2013).

To assess the effectiveness of the models, the complete dataset was resampled using a 'leave-group-out' cross-validation (Püschel et al., 2020). This method generates multiple splits of the data into modelling and prediction sets. This procedure was repeated 200 times and the data were divided into a modelling set containing 90% of randomly allocated observations, while the testing set contained the remaining 10%. The repetition number was selected to get stable estimates of performance and to mitigate overfitting. Overfitting means that the extracted model describes the behaviour of known data very well but does poorly on new data points (Pham & Triantaphyllou, 2007). By using this data splitting procedure, the model was not trained and then tested with the same dataset, instead it was tested the model's ability to predict new data that was not used in estimating it, in order to flag problems like overfitting.

Additionally, the seed was set at a random number to ensure reproducibility and 'increased' the sample to correct for class imbalance to standardize the classes. Different categories have different numbers of individuals per category, and if the class imbalance is severe, it might be a problem.

The *caret* package also provides a grid search (automatic and manual) where it is possible to specify tuning parameters for the models to work better. First, it started with an automatic grid search, setting the 'tune length option to indicate the number of different values to be tried for each algorithm parameter. In most of the tested models, this parameter was set to 10, in the random forest, it was set it to 4, because in this case, as there were 5 variables and they need to be divided into n number of trees, the maximum number of variables randomly sampled as candidates at each split here cannot be more than 4. This supports only integer algorithm parameters, thus providing a quick first attempt at guessing which values to try and which models are most promising. The best performing model can be further tuned using a more targeted/specific search.

The best classification models obtained for the ecomorphological data were then used to infer the main palaeoenvironment of the Pliocene fossil sample by computing their class probabilities to belong to each one of the Habitat and 'Wet or Dry' categories.

2.2.2) Ecometric analysis

To document the existence of an ecometric pattern, additionally to the abiotic conditions and measurement of functional characteristics of species, is also required the geographical ranges of species. And this is how it come to deal with functional characteristics at the community level.

Species occurrence and climatic conditions

Both types of data were gathered from online databases and available literature. Species range maps for existing species was obtained from the 'International Union for Conservation of Nature' Red List of Threatened Species website (IUCN, 2015). The occurrence data for existing species was also from Barr's (2014) published data, but instead of directly using that dataset, it was used the dataset from Barr (2017), where he compiled the species-mean values of the astragalus variables for each of 46 sub-Saharan African bovid species (detailed in Appendix A, Table A2). Barr (2017) removed some species just to included species for which the range map data were available. In addition, he also removed three species *Ammotragus lervia*, *Gazella cuvieri*, and *Addax nasomaculatus,* based on them either living alone, hence not being part of a wider community or living in the Sahara (one environment in a huge area). When using community trait values, care should be taken not to have one specimen dictating the average for an area, as this would influence the predictions. Whereas for fossil species is the same dataset that has been also applied to the astragalar ecomorphology.

Climate variables were collected from https://www.worldclim.org/. WorldClim databases provide relatively high-resolution datasets of climate for the globe (Hijmans et al., 2005). The latest version WorldClim version 2 (Fick & Hijmans, 2017) contains average monthly climatic gridded data for the period 1970-2000.

Bioclimatic variables are derived from the monthly temperature and rainfall values to generate more biologically meaningful variables. It was used Mean Annual Temperature (ºC * 10) and Mean Annual Precipitation (mm) to represent annual trends, Temperature Seasonality (standard deviation ×100) and Precipitation Seasonality (Coefficient of Variation) (%) to represent seasonality, and finally, Max Temperature of Warmest Month, Min Temperature of Coldest Month, Wettest Month Precipitation and Driest Month Precipitation to represent extreme or limiting environmental factors. Precipitation Seasonality captures dispersion in relative terms, it means that if there is a high variation of precipitation in that region, then the seasonality is high. Temperature seasonality is a measure of the variation in temperature over the year. The higher the standard deviation (the unit is temperature (ºC x10)), the greater the variability in temperature. The bioclimatic variables used can be seen in figure 8.

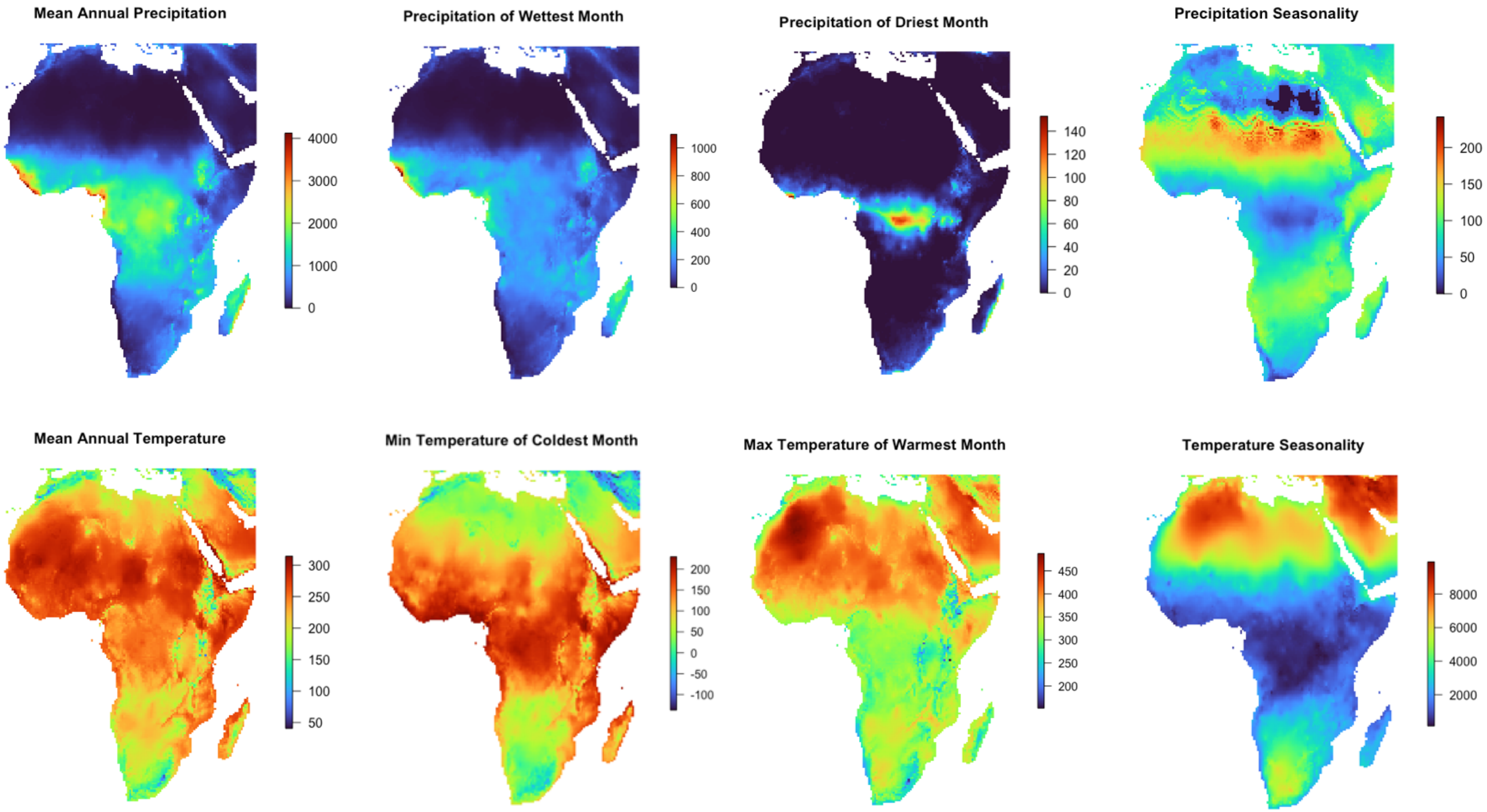

Figure 8 Bioclimatic variables used to test for trait–climate relationships. Mean annual precipitation and precipitation of wettest and driest month is in millimeters. Precipitation seasonality is a percentage. Mean annual temperature and extremes values of coldest and warmest month are in °C x 10. Temperature seasonality is in (°C x 10) x 100.

Land cover raster is from Broxton et al. (2014). The land cover variables were simplified to only include land cover classes well represented in sub-Saharan Africa (Table 2).

Table 2 Land cover classes with codes and descriptions following Friedl et al. (2010). Only land cover classes analysed in this study are included in the table.

Code	Description	Number of pixels in raster
2	Evergreen broadleaf forest	932
7	Open shrublands	906
8	Woody savannahs	1567
9	Savannahs	1705
10	Grasslands	860
11	Permanent wetland	59
16	Barren or sparsely vegetated	366

These land cover classifications are defined by the International Geosphere-Biosphere Programme (IGBP), detailed in Friedl et al. (2002):

Evergreen broadleaved forest is land dominated by broadleaved woody vegetation with a percentage cover >60% and height greater than 2 m. Almost all trees and shrubs remain green all year round. The canopy is never without green foliage. Open shrubland, on the other hand, is land with woody vegetation less than 2 m high and with shrub cover between 10% and 60%. The foliage of the shrubs can be evergreen or deciduous. Woody savannahs are characterised by land with herbaceous vegetation and other understory systems, and with forest canopy cover between 30% and 60%. The forest cover height exceeds 2 m, as in savannahs. However, in the case of the last, the forest cover is only between 10% and 30%. Finally, grasslands, which are also lands with herbaceous types of cover, have a tree and shrub cover of less than 10%. Permanent Wetland are characterised by land with a permanent mixture of water and herbaceous or woody vegetation. The vegetation may be present in either salt, brackish or fresh water. Barren or sparsely vegetated are lands with exposed soil, sand, rocks, or snow and never have more than 10% vegetation cover during any time of the year.

Data analysis

All spatial and statistical analyses were performed using RStudio Version 1.3.1073 (RStudio Team, 2020). The packages used were *rgdal* v.1.5-23. (Bivand et al., 2021), which provides bindings for the 'Geospatial' Data Abstraction Library; *rgeos* v.0.5-5. (Bivand &

Rundel, 2020), which serves as an interface to Geometry Engine - Open Source ('GEOS'); *maptools* v.1.1-1. (Bivand & Lewin-Koh, 2021), tools for handling spatial objects; *dplyr* v.1.0.5. (Wickham, et al., 2021), tool for working with data frames; *raster* v.3.4-5 (Hijmans, 2020), for geographic data analysis and modelling; *tidyr* v.1.1.3. (Wickham & Henry, 2020), tools to help to create tidy data; *reshape2* v.1.4.4 (Wickham, 2007), to restructure and aggregate data; *GISTools* v.0.7-4 (Brunsdon & Chen, 2014), mapping and spatial data manipulation tools; *viridis* v. 0.6.1 (Garnier, 2021), colour-blind-friendly colour maps for R; *cartography* v.3.0.0. (Giraud & Lambert, 2021), create and integrate maps in the R workflow; *caret* v. 6.0-88. (Kuhn, 2021), for classification and regression training; *CAST* v.0.5.1. (Hanna, 2021), 'caret' applications for spatial-temporal models; and *imager* v.0.42.8. (Barthelme, 2021), an image processing library.

First, all the data had to be merged. To achieve this, the map of Africa was transformed into a raster, which is an image with the surface divided into a regular grid of pixels. In this case, each grid cell was 50x50km. Next, using the bovid geographic distribution, it is determined which modern-day species intersected in each grid cell, based on their taxonomic name, and is merged with the African map, to mimic the computational locations. The same was done for the climatic condition’s variables. Hence, each grid cell represented a list of species occurrence and the associated climatic parameters, which are the bioclimatic variables and the land cover categories. The bioclimatic variables values were square root transformed to improve normality, just so that the measurements are in terms of the unit area of the grid/raster. However, the results will be presented in the original unit length.

Afterwards, for each grid cell, the species-mean values of each astragalus variable are computed. Thus, for each grid cell on the map, there is a mean value of each of the measurements. These mean values are then used as inputs for prediction models with the climate parameters as target variables.

Finally, a regression technique was used to test the relationship between the ecometric variables with the different bioclimatic variables in each grid. While for land cover, a classification technique was used.

There are some differences between the classification and regression problems. Fundamentally, classification is about predicting a discrete class label and regression is about predicting a continuous quantity (Kuhn & Johnson, 2013). Nevertheless, some algorithms can be used for both classification and regression (Kuhn & Johnson, 2013).

Model preparation for bioclimatic variables

The performance of the regression models was measured using the metrics Root-mean-square error (RMSE), mean absolute error (MAE) and coefficients of determination (R^2). The first two are the most commons metrics for regression models (Chai & Draxler, 2014). RMSE is interpreted as how far, on average, the residuals are from zero (Kuhn & Johnson, 2013). On the other hand, the calculation of MAE involves summing the magnitudes (absolute values) of the errors to obtain the 'total error' and then dividing the total error by n (Willmott & Matsuura, 2005). These metrics tell us how accurate our predictions are and, what is the amount of deviation from the actual values. Coefficients of determination (R^2) was introduced by Wright (1921), its original formulation quantifies how much the dependent variable is determined by the independent variables, in terms of proportion of variance. Regarding MAE and RMSE metrics, lower values indicate a better model. For R^2, the model closer to 1 is the best, since this metric only goes from 0 to 1.

To evaluate the model's execution, the complete dataset was resampled using the K-Fold cross-validation process with k = 10. As explained by Kuhn and Johnson (2013), in the K-Fold Cross-Validation the samples are randomly partitioned into k sets of roughly equal size. A model is fit using all samples except the first subset (called the first fold). The held-out samples are predicted by this model and used to estimate performance measures. The first subset is returned to the training set and the procedure repeats with the second subset held out, and so on. The k resampled estimates of performance are summarized (usually with the mean and standard error) and used to understand the relationship between the tuning parameter(s) and model utility (Kuhn & Johnson, 2013). As the sample size of the ecometric dataset is particularly large, the leave-out group would have taken a long time to process.

All data was centred and scaled as a pre-process to improve the numerical stability of some subsequent calculations and standardize their scale. As a result of centring, the variables have a zero mean, while scaling coerce the predictors to have a common standard deviation of one. These transformed interval values were, subsequently, used in the analyses. Additionally, the seed was set at a random number to ensure reproducibility.

Then four different regression models (supervised algorithms) were executed, (i) linear model; (ii) k-nearest neighbours (KNN); (iv) support vector machine (SVM) and (v) random forest (RF). The last three have already been described above, the only difference is that they will now be used for regression problems.

(i) linear model

Linear regression is a linear model wherein a model assumes a linear relationship between the input variables (x) and the single output variable (y) (Yan & Su, 2009). When there is a single input variable (x), the method is called simple linear regression. When there are multiple input variables, the procedure is referred to as multiple linear regression.

Again, most of the models tested had the tunning parameter set to 10, and in the random forest was set to 4.

Model preparation for land cover

For the land cover, the process is similar to ecomorphology as it also deals with categories since it is a discrete variable. Thus, four classifiers (supervised algorithms) were chosen: (i) linear discriminant analysis (LDA); (ii) k-nearest neighbours (KNN); (iv) support vector machine (SVM) and (vi) random forest (RF). They have already been described above in the ecomorphological analysis.

The performance of the classification models was quantified using the confusion matrix from which the overall classification accuracy (i.e., error rate) was computed, as well as by computing Cohen’s Kappa coefficient (Püschel et al., 2018).

To assess the model's effectiveness, the complete dataset was resampled using the K-Fold cross-validation process with k = 10.

The tuning parameters are the same as in the previous analyses.

3) Results

3.1) Exploratory analysis

Principal component analysis (PCA)

Two fossil specimens from Kanapoi and Allia Bay were removed from the database because they corresponded to outliers, probably representing measurement errors. Then, a Principal component analysis (PCA) was carried out again (Figure 9).

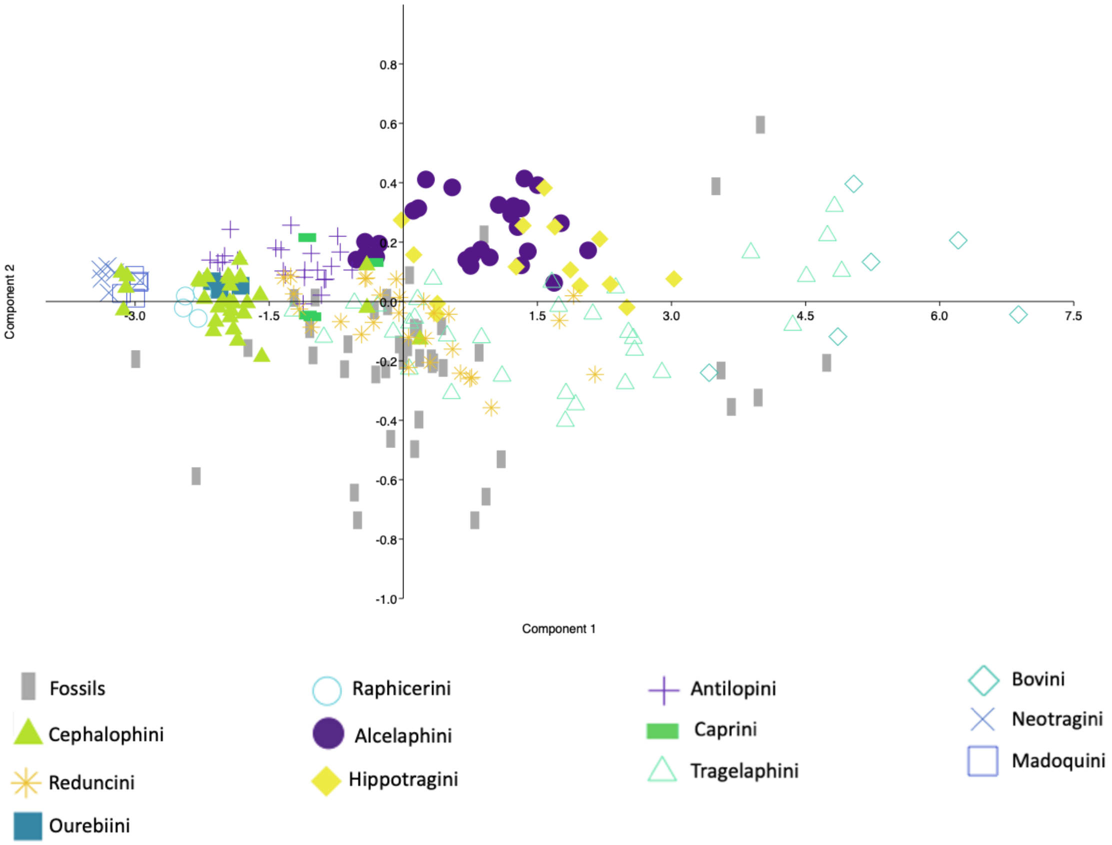

Figure 9 Principal component analysis (PCA) of the astragalar measurements (only the first two PCs are shown). The models located at the extremes of the PC axes correspond to the individuals showing the maximum or minimum values along the first two PCs. Data is coloured and shaped by the tribe name.

The principal components analysis of the astragalar measurements shows a high covariation among the five variables. The first principal component (PC) vector explains 98.46% variance in the data (eigenvalue = 3.95), whereas the other four PCs explained 1.07%

(eigenvalue = 0.042), 0.25% (eigenvalue = 0.010), 0.13% (eigenvalue = 0.005) and 0.07% (eigenvalue = 0.003), respectively.

The first PC explains almost 100% of the variance, which means that almost all the variance represented by the five original variables can be effectively summarized by the linear transformation represented by the first PC, which is an extreme case of linear dependence. This occurs because the variables in this study are highly collinear, which means that probably just one or two of the original variables could effectively describe the observed patterns.

This is a common issue when dealing with traditional morphometric data (i.e., linear distances), as size is a confounding factor that seems to be explaining almost all the variance of the sample. Thus, the astragalar data was corrected for size and the following figures (Figure 10 a-c) show a PCA of these transformations, as well as of the raw data for comparison purposes.

With data scaled by geometric mean or transformed by Burnaby's method, the values for the different PCs are more evenly distributed. For the Burnaby isometric transformation, the first PC explained 59.06% and the second 23.3%, and for the geometric mean transformation data, the first PC explained 65.53% and the second 17.75%. Thus, the first PC no longer explains almost all the variation.

However, when removing the effect of size, there is a loss of differentiation between the different habitats. It is only in the scatter plot of the first two PC axes of the raw data that there is a slight pattern of differentiation between animals occupying different habitats and yet most of the different habitats occupy both negative and positive values in the first PC. Thus, it could be the size of the animals explaining the distinction between habitats rather than the astragalus measurements. Having said that, the following analyses were done using different approaches commonly applied to deal with collinear data, such as using size-standardized data, as well as reducing the number of variables, among other options. The data with the best results was the one that did not go through any transformation, i.e., raw data. The analyses were also repeated using fewer variables, but the results were worse than those discussed below and are not presented.

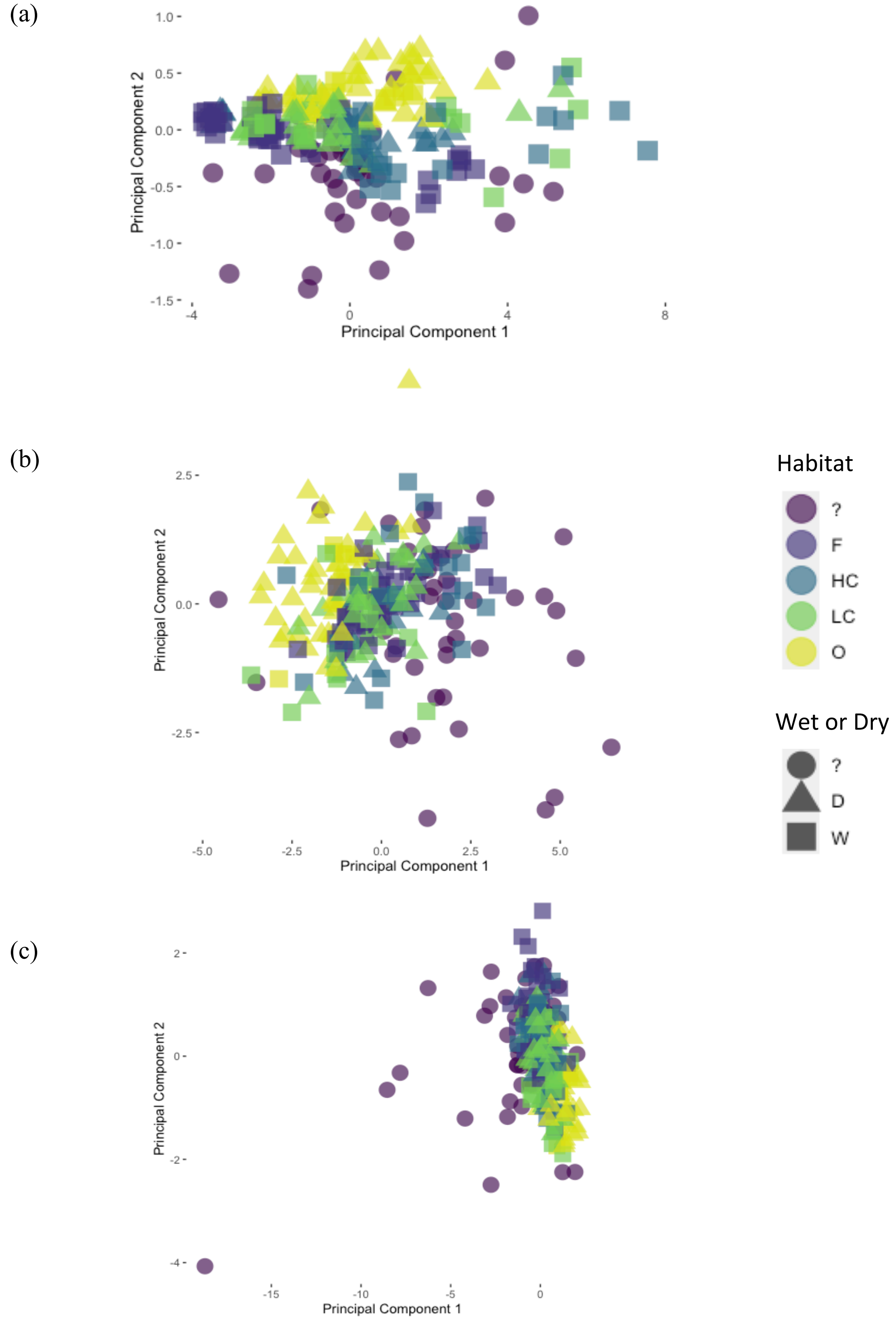

Figure 10 Principal component analysis (PCA) of the astragalar measurements (only the two first PCs are shown). (a) Using the raw data, (b) using the Isometric transformation data, and (c) using the data scaled by geometric mean. The models located at the extremes of the PC axes correspond to the individuals showing the maximum or minimum values along the two first PCs. Data is coloured based on habitat type and shaped based on 'wet and dry' categories.

3.2) Ecomorphological analysis

Habitat Classification

The statistical results are presented in Table 3, which shows the accuracy and Cohen's Kappa results for all applied models after performing the leave-group-out cross-validation procedure. The most accurate model for predicting habitat was Random Forest (RF).

To all models tested it was applied an automatic grid search, and RF model performed better. Hence, the obtained RF model was further tuned using a manual grid search. Two parameters were tuned in this model, the number of trees to grow (i.e. 100, 200, 500, 1000 and 2000) as well as the number of variables randomly sampled as candidates at each split (i.e. 1, 2, 3 and 4). In general, the RF model was quite robust when changing these tuning parameters, showing similar classification accuracies. The final best RF model grew 2000 trees and used four variables randomly sampled as candidates at each split (average accuracy: 0.712; average Cohen's Kappa: 0.613) (Table 3).

Table 3 Summary statistics of the "leave-group-out" cross-validation procedure for the accuracy and Cohen's kappa values for all the tested models (best results in bold).

	RF	SVM	LDA	CART	KNN	NB
Accuracy	**0.7118750**	0.6778125	0.6371875	0.6321875	0.6103125	0.5368750
Kappa	**0.6134392**	0.5665606	0.5118211	0.5096735	0.4796661	0.3875724

The random forest model is statistically different from all other models (p-value <0.05), with the highest difference values, both in accuracy and Cohen's kappa.

To allow a visual representation of the separation between habitat categories, the RF model was used to plot a Decision Boundary (Figure 11a). The model used the two variables that most contributed to the separation between categories and led to the identification of a linear pattern, allowing some discrimination between ungulate species inhabiting open and mixed/closed environments. Figure 11b shows the ecomorphic variables sorted by importance, with radius of the distal articular end (Distrad) and Functional length (B) having the most contribution, while the remaining variables explain less variation.

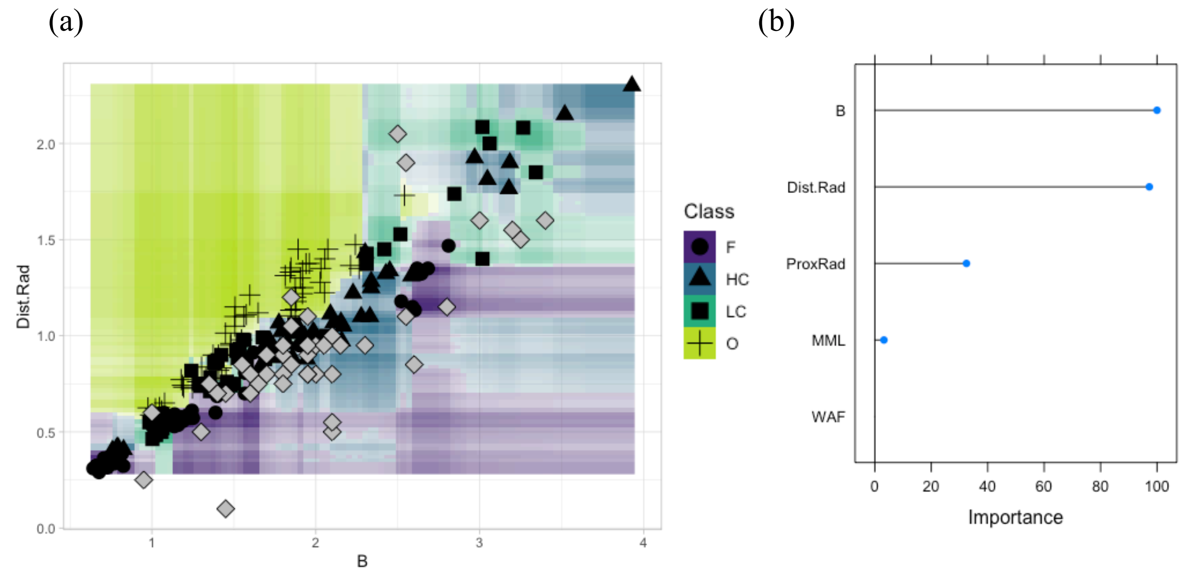

Figure 11 (a) Decision boundary plot. The two variables that contributed the most to the RF model are displayed. The space is colored depending on what habitat preference the RF algorithm predict that region belongs to, whereas the lines between colored areas represent the decision boundaries. Color intensity indicates the certainty of the prediction in a particular graph area (i.e., darker colors mean a higher probability of belonging to a certain category). The grey rhombuses correspond to the fossil specimens. In addition, (b) variable importance scores for habitat predictors used are provided.

The bivariate plot in Fig. 11a shows that, with the only exception of species ecological adapted to open environments, it was not possible to make a clear distinction between species from mixed environments (light cover and heavy cover) or those belonging to closed environments.

By applying the final RF model, the fossil sample was classified, and all the specimens were categorized as living in Open, Light Cover, Heavy Cover or Forest habitat, as presented in Table 4.

Table 4 Random Forest prediction classification of habitat category for the fossil sample (Kanapoi and Allia Bay). Detailed results for individual fossils are provided in Appendix C, Table C1.

Assemblage	**Open (n, %)**	**Light cover (n, %)**	**Heavy cover (n, %)**	**Forest (n, %)**	**Total**
Allia Bay	1 (7%)	5 (36%)	5 (36%)	3 (21%)	14 (100%)
Kanapoi	4 (13%)	12 (40%)	9 (30%)	5 (17%)	30 (100%)

'Wet & Dry' Classification

The Random Forest (RF) model was again found to be the most accurate, this time for wet or dry prediction. The statistical results are presented in Table 5, which shows the accuracy and Cohen's Kappa results for all applied models after performing the leave-group-out cross-validation procedure.

Again, an automatic grid search was applied to all models tested, and the RF model showed a better performance. Thus, as in the previous analysis the RF model obtained was tuned using exactly the same parameters. In these adjustments, the RF model was quite robust, showing similar classification accuracies. The best final RF model grew 2000 trees and used four randomly sampled variables as candidates in each split (average accuracy: 0.755; average Cohen's Kappa: 0.507) (Table 5).

Table 5 Summary statistics of the "leave-group-out" cross-validation procedure for the accuracy and Cohen's kappa values for all the tested models (best results in bold).

	RF	**SVM**	**LDA**	**CART**	**KNN**	**NB**
Accuracy	**0.7550000**	0.7111765	0.7044118	0.7008824	0.7220588	0.6326471
Kappa	**0.5070269**	0.4176269	0.4047612	0.3961819	0.4401833	0.2508601

The random forest model is statistically different from all other models (p-value <0.05), with the highest difference values, both in accuracy and Cohen's kappa.

Plotting the two variables that contribute most to the RF model led to the identification of a linear pattern that allows a fairly good discrimination between bovid species belonging to wetter or drier habitats (Figure 12a). Of the five variables used, the Functional length (B) and the radius of the distal articular end (Distrad) have the most importance and contributed the most to the separation between categories, while the remaining variables explain less variation (Figure 12b).

Compared to the previous bivariate plot (Fig. 11a), this plot shows that there is an association between the Dry category and the Open Habitat category, and between the Wet category and the other categories (light and heavy cover and forest) (Fig. 12a).

By applying the final RF model, the fossil sample was classified, and all the specimens were categorized as living in Dry or Wet habitats (Table 6).

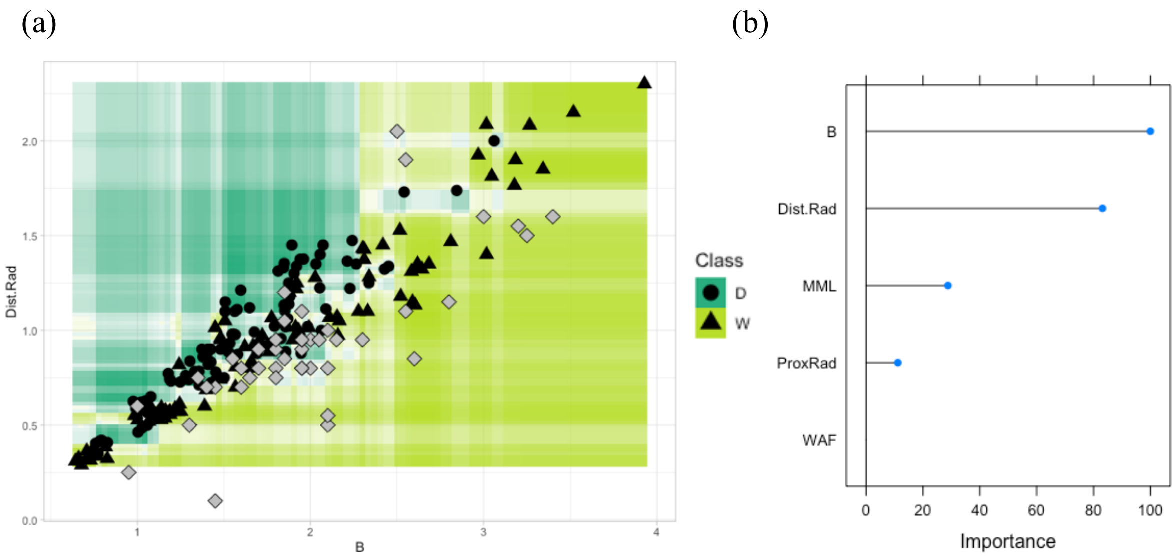

Figure 12 (a) Decision boundary plot. The two variables that contributed the most to the RF model are displayed. The space is colored depending on what habitat preference the RF algorithm predict that region belongs to, whereas the lines between colored areas represent the decision boundaries. Color intensity indicates the certainty of the prediction in a particular graph area (i.e., darker colors mean a higher probability of belonging to a certain category). The grey rhombuses correspond to the fossil specimens. In addition, (b) variable importance scores for habitat predictors used are provided.

Table 6 Random Forest prediction classification of 'Wet and Dry' category for the fossil sample (Kanapoi and Allia Bay). Detailed results for individual fossils are provided in Appendix C, Table C2.

Assemblage	**Wet (n, %)**	**Dry (n, %)**	**Total**
Allia Bay	6 (43%)	8 (57%)	14 (100%)
Kanapoi	19 (63%)	11 (37%)	30 (100%)

3.3) Ecometric analyses

The geographical distribution, in Africa, of the modern species used in the comparative data is illustrated in Figure 13a. In Figure 13b, they are represented by their richness (i.e., the number of modern species intercepting in each grid cell). This is simply a count of the number of species used and it does not consider the abundances of the species or their relative abundance distributions. As can be observed in figure 13b, the highest species richness (i.e., number of species in a certain area) occurs in East Africa.

Figure 14 illustrates the spatial distributions of modern-day species community-mean values of each of the five astragalus variables.

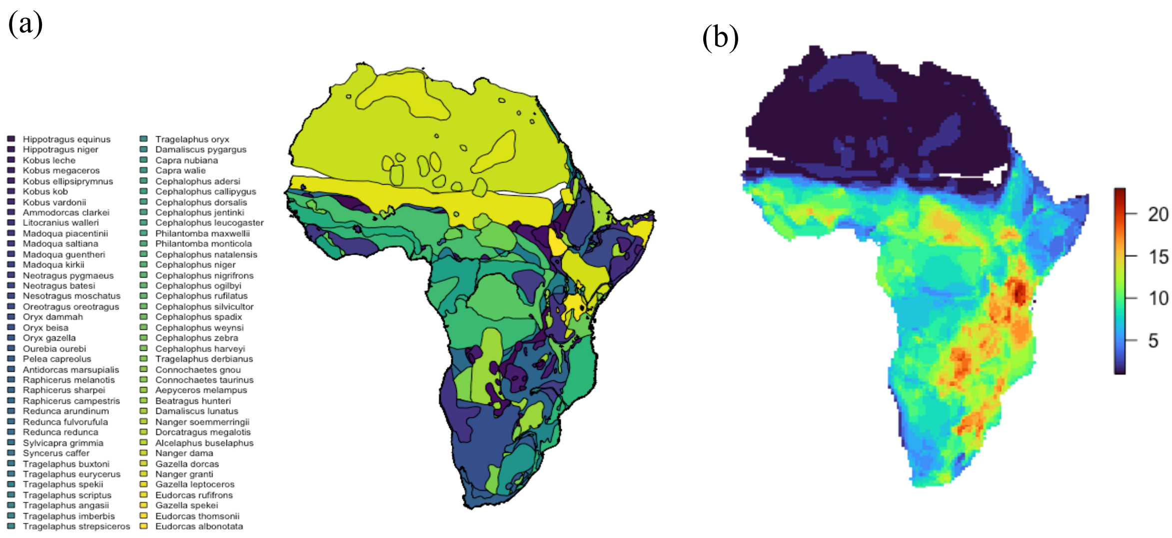

Figure 13 Map of the African continent showing (a) the geographical distribution of modern-day species used in this study and (b) species richness.

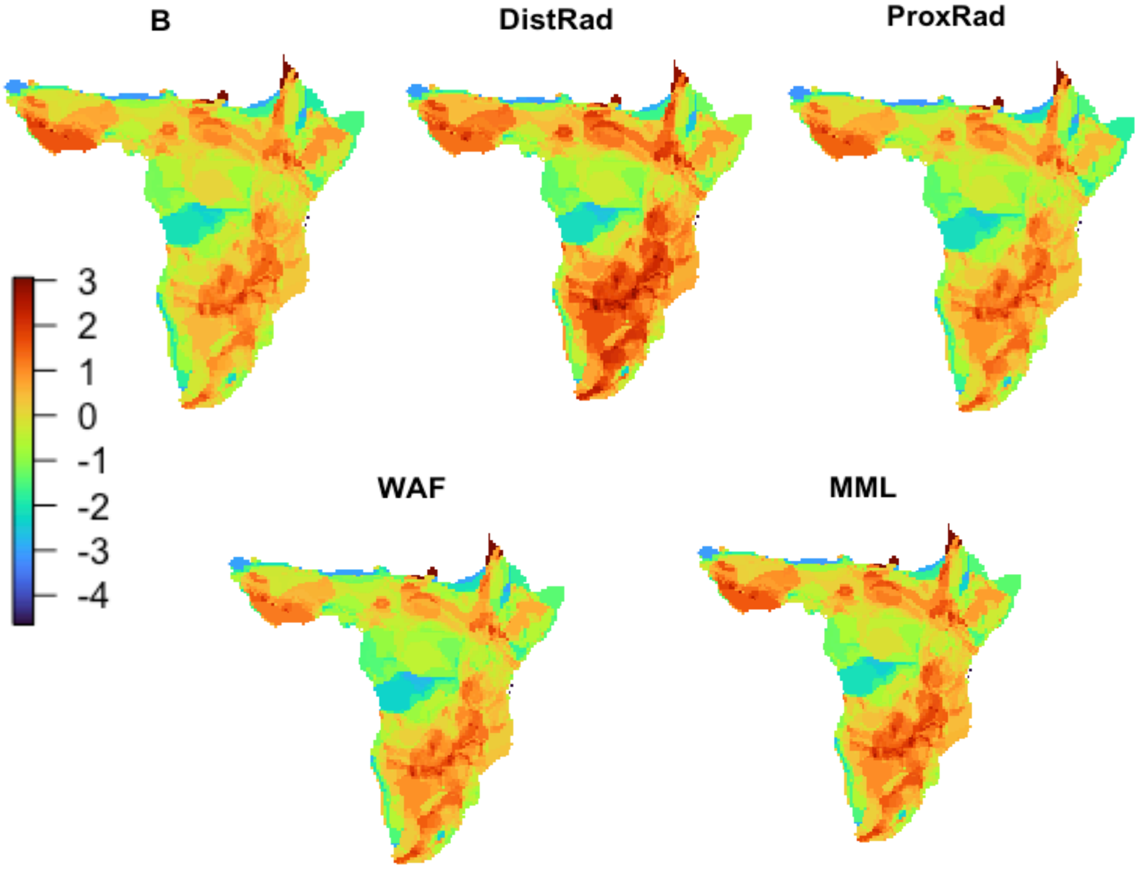

Figure 14 Rasters showing the spatial distribution of assemblage-average ecometric values for the five variables. Variables acronyms: Functional length (B); Radius of circle fit to the margin of the distal trochlea (DistRad); Radius of circle fit to the margin of the proximal trochlea (ProxRad); Width at level of flange on lateral surface, but excluding flange (WAF); Maximum medial length (MML).

Ecometric precipitation estimates

On Table 7 are presented the results of MAE, RMSE and R^2, after performing the k-fold cross-validation procedure, for all models applied to each bioclimatic variable related to precipitation (i.e., mean annual precipitation, precipitation of the wettest and driest month and precipitation seasonality). The best model, i.e., the one with the lowest error (MAE and RMSE) and the best coefficient of determination (R^2) to predict all the bioclimatic variables was Random Forest (RF).

An automatic grid search was applied to all the tested models, and the RF model had the best performance in all bioclimatic variables tested here. Therefore, each RF model obtained for each bioclimatic variable was further tunned using a manual grid search. Similar to the RF model used in the ecomorphological analysis, the same two parameters were tuned in this model, the number of trees to grow (i.e. 100, 200, 600, 1000 and 2000), as well as the number of variables randomly sampled as candidates in each division (i.e. 1, 2, 3 and 4), although in this case it is a regression model rather than a classification model.

In general, all RF models were quite robust when changing these tuning parameters, showing similar predictions for all variables. In the end, all these RF models grew 600 trees and used three randomly sampled variables as candidates in each split. The results for each model can be seen in the table below (Table 7).

Table 7 Summary statistics of the "K-Fold" cross-validation procedure for the MAE, RMSE and R^2 mean values for all the tested models (best results in bold).

Precipitation estimates	**Model**	**Accuracy**		
		MAE	RMSE	R^2
Mean Annual Precipitation	RF	**2.174750**	**3.247268**	**0.8749964**
	SVM	3.053304	4.303970	0.7810089
	KNN	2.563698	3.871536	0.8234463
	LM	4.451616	5.911415	0.5858922
Precipitation of Wettest Month	RF	**0.9858932**	**1.530273**	**0.8083321**
	SVM	1.4171792	2.005243	0.6714070
	KNN	1.1660399	1.785131	0.7408373
	LM	1.9715621	2.576234	0.4562863
Precipitation of Driest Month	RF	**0.6610180**	**1.019041**	**0.8148615**
	SVM	1.0092131	1.461645	0.6275441
	KNN	0.7654543	1.179462	0.7537227
	LM	1.5074221	1.944205	0.3276453
Precipitation Seasonality	RF	**0.5100571**	**0.7723129**	**0.7869078**
	SVM	0.7661778	1.0994718	0.5710157
	KNN	0.6093220	0.9179234	0.7011483
	LM	1.1252165	1.4687855	0.2297147

All models tested for each bioclimatic variable are statistically different (p-value <0.05), with RF having the highest difference values in all of them, both in MAE, RMSE and R^2.

By applying the final RF model of each bioclimatic variable, it is possible to create a spatial prediction of these precipitation variables using the astragalar variables from the comparative data. To achieve this, the RF model was used to predict associations of the mean of each ecometric variable with the values of the bioclimatic variable being used, in each grid cell on the map. Thus, for each bovid community (i.e., a 50x50km grid cell), particular astragalar measurements are associated with specific precipitation values.

Figure 15 presents the maps of the currant values, for Africa, taken from worldclim.com, and their estimates, across sub-Saharan Africa, for (a) mean annual precipitation, (b) precipitation of wettest month, (c) precipitation of the driest month and (d) precipitation seasonality. As well as the ecomorphic variables that contributed most to these predictions (Figure 15 1-4).

In general, of the five ecomorphic variables the radius of the distal articular end (Distrad) and Functional length (B) have the most importance and explain most of the variation in precipitation estimates, while the remaining variables explain less variation depending on the bioclimatic variable tested (Figure 15 1-4).

The fossil sample was then used to predict the values for Kanapoi and Allia Bay. The predicted values of the RF model for each of the precipitation bioclimatic variable are presented in Table 8.

Table 8 Random Forest prediction of Mean Annual Precipitation, Precipitation of Wettest Month, Precipitation of Driest Month and Precipitation Seasonality, for the fossil sample (Kanapoi and Allia Bay).

Assemblage	**Precipitation estimates**			
	Mean Annual Precipitation (mm)	Precipitation of Wettest Month (mm)	Precipitation of Driest Month (mm)	Precipitation Seasonality (%)
Allia Bay	853.2	198	25.3	58.2
Kanapoi	1701.9	281.8	27.7	60.1

These values were not predicted based on location but rather based on the average of the astragalus variables measurements of the fossils collected at these sites.

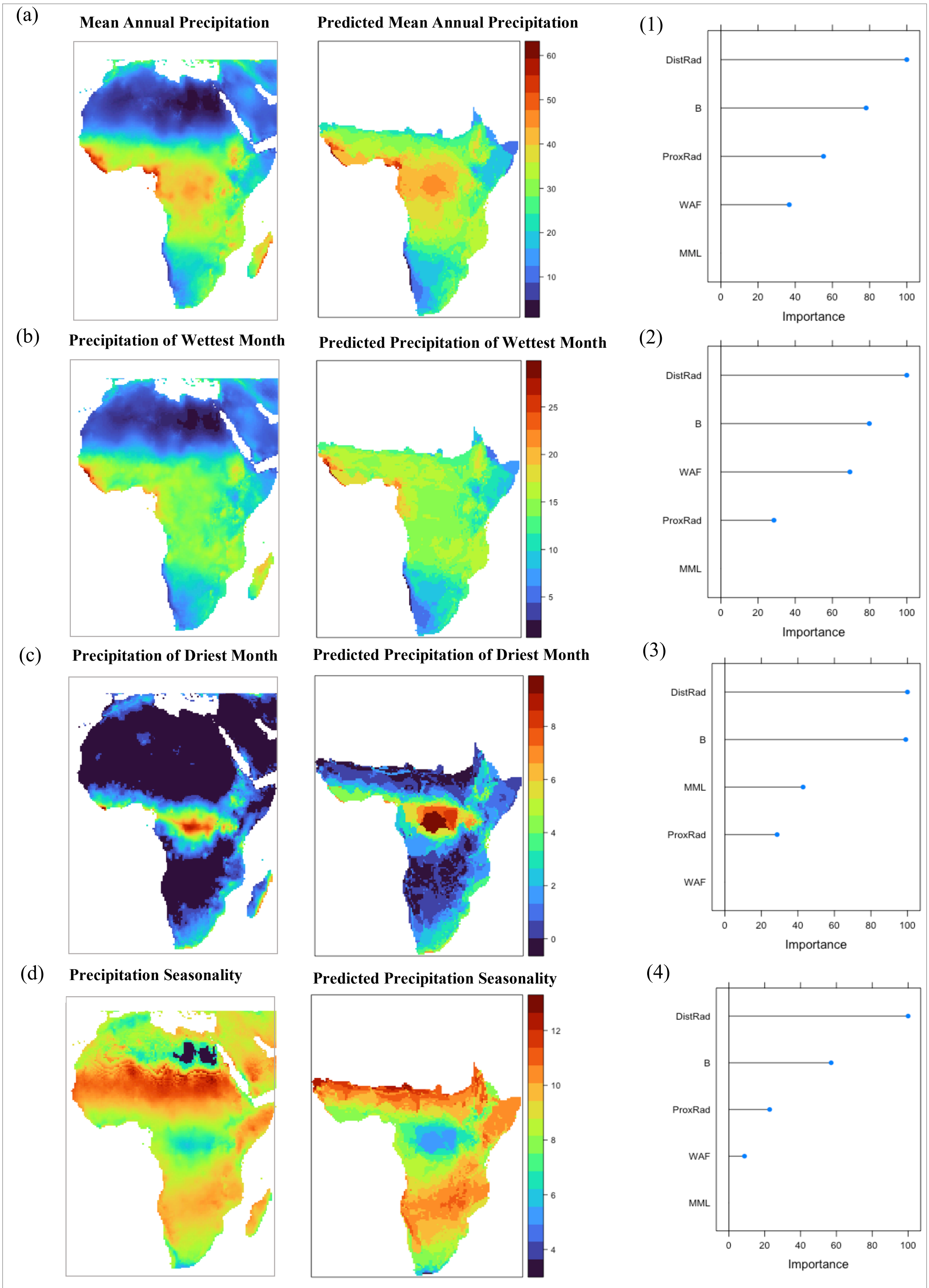

Figure 15 Maps showing precipitation estimates for Africa. For each precipitation variable tested, (a) Mean annual precipitation (MAP), (b) precipitation of wettest month (PWM), (c) precipitation of driest month (PDM) and (d) Precipitation seasonality (PS), there is a visualisation of the current values used for modelling (left) and their spatial prediction for sub-Saharan Africa (right), based on modern species ecomorphic variables using the RF model. Units are square root transformed. In addition, variable importance scores are provided for the predictors used for each of these bioclimatic variables, i.e. (1) for MAP, (2) for PWM, (3) for PDM and (4) for PS.

Ecometric temperature estimates

The results of MAE, RMSE and R^2, after performing the k-fold cross-validation procedure, for all models applied to each bioclimatic variable related to temperature (i.e., mean annual temperature, max temperature of warmest month, min temperature of coldest month and temperature seasonality), are presented in Table 9. The best model, i.e., the one with the lowest error (MAE and RMSE) and the best coefficient of determination (R^2) to predict all the bioclimatic variables was again Random Forest (RF).

An automatic grid search was applied to all the tested models, and the RF model had the best performance in all bioclimatic variables tested here. Therefore, each RF model obtained was further tunned using a manual grid search, the exact same two parameters were adjusted in this RF model.

In general, all RF models were quite robust when changing these tuning parameters, showing similar predictions for all variables. In the end, all these RF models grew 600 trees and mean annual temperature used three randomly sampled variables as candidates in each split, while the rest of the bioclimate variables used two variables. The results for each model can be seen in the table below (Table 9).

Table 9 Summary statistics of the "K-Fold" cross-validation procedure for the MAE, RMSE and R^2 mean values for all the tested models (best results in bold).

Temperature estimates	**Model**	**Accuracy**		
		MAE	RMSE	R^2
Mean Annual Temperature	RF	**0.3791258**	**0.6075898**	**0.7312159**
	SVM	0.5388479	0.8057513	0.5374535
	KNN	0.4447440	0.7119740	0.6337903
	LM	0.7250144	0.9748980	0.3091169
Max Temperature of Warmest Month	RF	**0.3688247**	**0.5737843**	**0.7004779**
	SVM	0.5278867	0.7663771	0.4685141
	KNN	0.4297809	0.6599279	0.6076861
	LM	0.7322884	0.9648266	0.1526834
Min Temperature of Coldest Month	RF	**0.7381527**	**1.129270**	**0.8317325**
	SVM	1.0897428	1.542461	0.6879371
	KNN	0.8732290	1.333771	0.7663766
	LM	1.4682730	1.903016	0.5200701
Temperature Seasonality	RF	**3.333054**	**4.956625**	**0.8711849**
	SVM	5.153804	7.028412	0.7416879
	KNN	4.012857	5.967009	0.8142574
	LM	6.869985	8.952767	0.5798703

All models tested for each temperature estimate are statistically different (p-value <0.05), with RF having the highest difference values, both in MAE, RMSE and R^2.

By applying the final RF model of each bioclimatic variable, it is possible to create a spatial prediction of these precipitation variables using the astragalar variables from the comparative data.

Figure 16 presents the maps of the currant values, for Africa, taken from worldclim.com, and their estimates, across sub-Saharan Africa, for (a) mean annual temperature, (b) max temperature of warmest month, (c) min temperature of coldest month and (d) temperature seasonality. As well as the ecomorphic variables that contributed most to these predictions (Figure 16 1-4).

In general, of the five ecomorphic variables the Functional length (B), Width at level of flange on lateral surface (WAF) and the radius of the distal articular end (Distrad) have the most importance and explain most of the variation in temperature estimates, while the remaining variables explain less variation depending on the bioclimatic variable tested (Figure 15 1-4).

The fossil sample was then used to predict the values for Kanapoi and Allia Bay. The predicted values of the RF model for each of the temperature bioclimatic variable are presented in Table 10.

Table 10 Random Forest prediction of Mean Annual Temperature, Max Temperature of Warmest Month, Min Temperature of Coldest Month and Temperature Seasonality, for the fossil sample (Kanapoi and Allia Bay).

Assemblage	**Temperature estimates**			
	Mean Annual Temperature (ºC)	Max Temperature of Warmest Month (ºC)	Min Temperature of Coldest Month (ºC)	Temperature Seasonality (ºC)
Allia Bay	23.79	32.64	14.07	1.176
Kanapoi	25.7	33.64	16.12	1.073

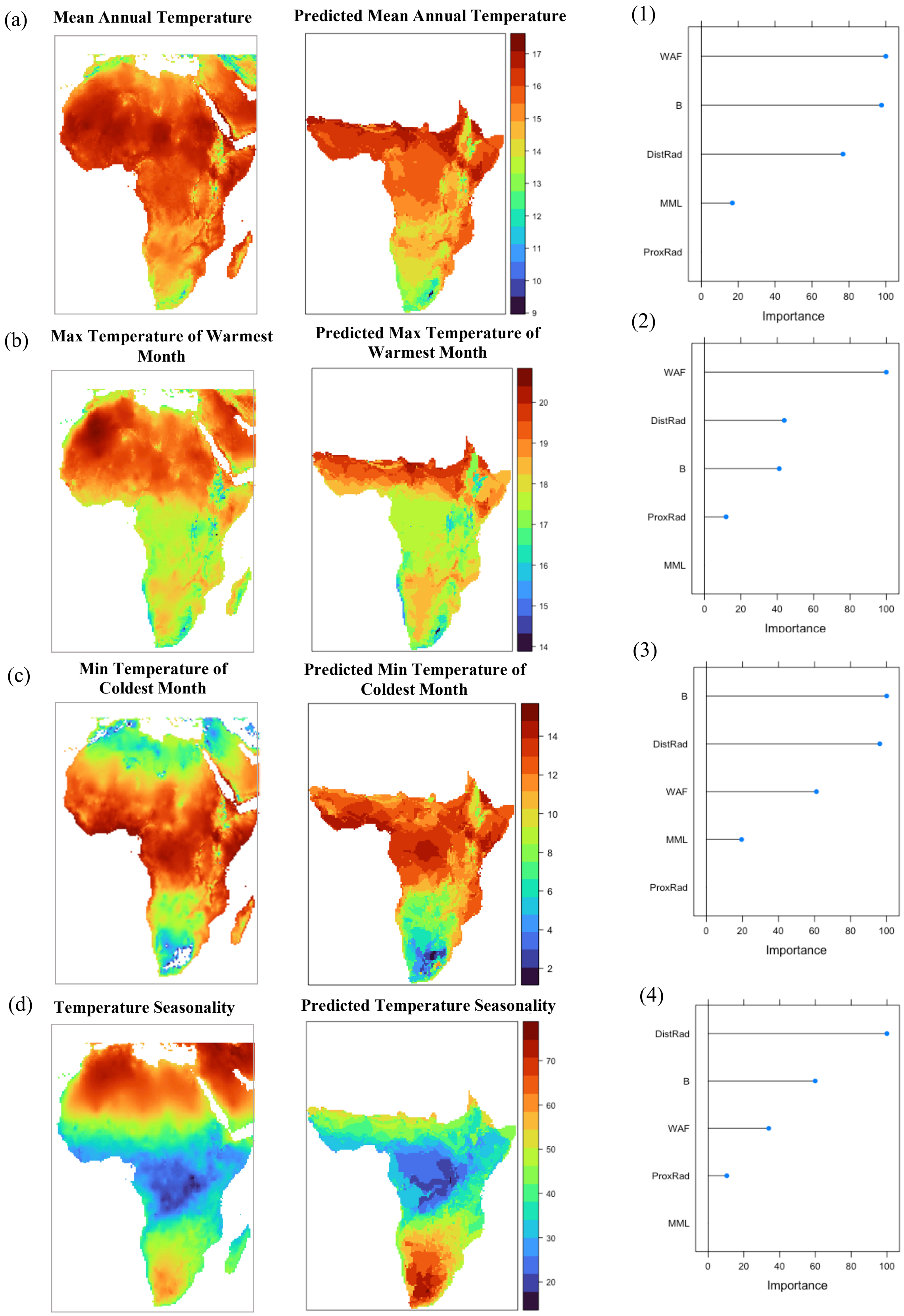

Figure 16 Maps showing temperature estimates for Africa. For each precipitation variable tested, (a) Mean annual temperature (MAT), (b) Max Temperature of Warmest Month (MTWM), (c) Min Temperature of Coldest Month (MTCM) and (d) Temperature Seasonality (TS), there is a visualisation of the current values used for modelling (left) and their spatial prediction for sub-Saharan Africa (right), based on modern species ecomorphic variables using the RF model. Units are square root transformed. In addition, variable importance scores are provided for the predictors used for each of these bioclimatic variables, i.e. (1) for MAT, (2) for MTWM, (3) for MTCM and (4) for TS.

Ecometric land cover estimates

The statistical results are presented in Table 11, which shows the accuracy and Cohen's Kappa results for all applied models after performing the k-fold cross-validation procedure. The most accurate model for predicting land cover was Random Forest (RF).

To all models tested it was applied an automatic grid search, and RF model performed better. Hence, the obtained RF model was further tuned using a manual grid search. Two parameters were tuned in this model, the number of trees to grow (i.e. 100, 200, 500, 1000 and 2000) as well as the number of variables randomly sampled as candidates at each split (i.e. 1, 2, 3 and 4). In general, the RF model was quite robust when changing these tuning parameters, showing similar classification accuracies. The final best RF model grew 600 trees and used one variable randomly sampled as candidates at each split (Accuracy: 0.7469232; Kappa: 0.6787663) (Table 11).

Table 11 Summary statistics of the "K-Fold" cross-validation procedure for Accuracy and Cohen's Kappa mean values for all the tested models (best results in bold).

	RF	SVM	KNN	LDA
Accuracy	**0.7469232**	0.5617654	0.7109609	0.6078916
Kappa	**0.6787663**	0.4374237	0.6325701	0.5002338

All models are statistically different (p-value <0.05), with RF having the highest difference values, both in Accuracy and Cohen's Kappa.

By applying the final RF model, it's possible to create a spatial prediction of the different land covers categories using the astragalar variables of the comparative data (Figure 17a). Among the five ecomorphic variables the Functional length (B) and radius of the distal articular end (Distrad) have the most importance and explain most of the variation in land cover estimates, while the remaining variables explain less variation (Figure 17b).

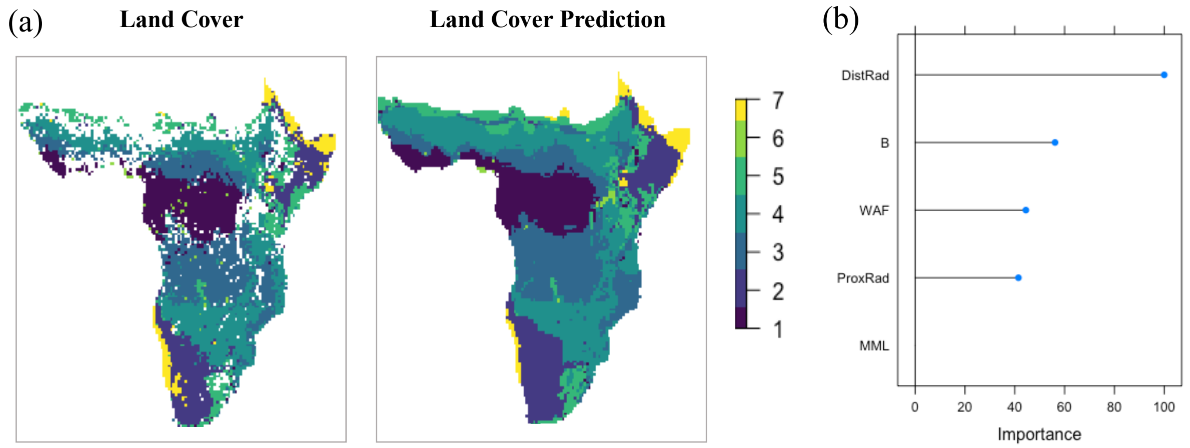

Figure 17 (a) Maps showing land cover for Africa. Visualisation of the modern-day land cover data used for modelling (left). Spatial prediction of sub-Saharan Africa land cover based on modern species ecomorphic variables using the RF model (right). Numbers correspond to: (1) Evergreen broadleaf forest, (2) Open shrublands; (3) Woody savannahs; (4) Savannahs; (5) Grasslands; (6) Permanent wetland; (7) Barren or sparsely vegetated. In addition, (b) variable importance scores for habitat predictors used are provided.

The fossil sample was then used to classify the data of Kanapoi and Allia Bay into different land covers categories (Table 12).

Table 12 Random Forest prediction classification of land cover for the fossil community sample (Kanapoi and Allia Bay).

Assemblage	**Land cover categories**						
	Evergreen broadleaf forest	Open shrublands	Woody savannahs	Savannahs	Grasslands	Permanent wetland	Barren
Allia Bay	0.260	0.048	0.172	0.078	0.380	0.052	0.010
Kanapoi	0.222	0.126	0.274	0.034	0.178	0.132	0.034

4) Discussion

Palaeoenvironmental reconstructions provide fundamental ecological contexts for key episodes in human evolution. Through ecomorphological and ecometric analyses, using the functional morphology of astragalus and testing several machine learning methods, it was possible to accurately predict the habitat of modern African bovid species and ultimately use these models to predict the palaeoenvironmental conditions of two important Early Pliocene fossil sites in East Africa. Several environmental variables were tested, with promising results, providing predicted values of averages, variations and extremes of temperature and precipitation variables, in addition to vegetation categories. This gives a robust idea of what sites related to *Australopithecus anamensis*, i.e., Kanapoi and Allia Bay would have been like. These results contribute towards the study of the environmental circumstances in which a key part of human history took place.

An initial exploratory analysis using PCAs showed that body size was highly correlated with most astragalar metrics, and that most habitat categories were not easily distinguishable. As previously mentioned by Barr (2014), this happens as the allometric effects of size (shorter and wider astragalus) are similar to the predicted functional effects of habitat-specific locomotor adaptation. Consequently, there is some overlap amongst different values, making it difficult to distinguish between different categories. Such overlap is a direct consequence of body size as a covariance factor, considering the way the astragalus adapts is also being influenced by changes occurring in other regions of an animal's morphology (Püschel et al., 2018). After all, body size is one of the most fundamental biological characteristics of mammals and has been shown to correlate significantly with many biological variables (Kovarovic et al., 2018). Nonetheless, it was still possible to detect habitat effects independent of body size using size-standardized data. However, it was determined that size correction was not beneficial for the ecomorphological analysis, as size seem to be an important factor when classifying astragali based on habitat characteristics. Therefore, the raw data was preferred as it resulted in a more accurate prediction of the environment. However, it is also likely that the use of more sophisticated tools to quantify morphology (e.g., geometric morphometrics [GM], elliptic Fourier analysis, Euclidean distance matrix analysis, among other options) would improve the performance of the models, as shape could be captured in a more precise manner (Benítez & Püschel, 2014).

GM is rapidly becoming the standard method in biological and palaeontological studies focused on morphology (Curran, 2018). Rather than analysing linear measurements, 3D GM quantifies and analyses morphology using a series of cartesian coordinates of an anatomical region of interest (Lawing & Polly, 2010). By using an algorithmic procedure known as the generalised Procrustes superimposition, it allows for a mathematically consistent separation of shape and size (Bookstein, 1991). Thus, it determines changes in the shape of objects independent of size effects and without a priori assumptions about which features are significant (Zelditch et al., 2004).

Ecomorphological results

Astragalar ecomorphology results show fossil bovids linked to each of the four habitat types, defined a priori, with better representation in the intermediate categories, i.e., light and heavy cover. These two categories combined represent about 70% of the predicted habitat for Allia Bay and Kanapoi, including habitats with bushland and light bushland, tall-grass habitats, woodland, as well as swampy habitats.

For the 'wet and dry' categories, the result for Allia Bay showed that most fossils were classified within the 'dry' category (57%), whereas most of Kanapoi's specimens were classified as 'wet' (63%). This means that Allia Bay has slightly more individuals representing a dry substrate. Kanapoi, on the contrary, has slightly more individuals representing a wetter substrate. However, it cannot be determined to be a drier or wetter habitat as the results are very close to 50%. Instead, Kanapoi and Allia Bay, probably represent a mix of the 'wet and dry' categories. Overall, these two sites do not differ much, at least when compared using astragalar bovid ecomorphological proxies.

Through the decision boundary plot, it was observed that the RF model was unable to establish a good separation between light and heavy cover and forest categories. Hence, it could mean that the specimen astragalus belonging to these categories are not very good at distinguishing between different complex habitats, only between open and complex habitats. After all, this type of environment already integrates some complexity in the vegetation structure and the main distinction between these categories is the density of vegetation, which creates obstacles around (or below) which animals must move to escape predators (Curran, 2012). Furthermore, when dealing with ecological and biological data that cannot be easily structured as discrete units, such as habitats and skeletal morphologies of closely related species, some overlap between groups is always expected (Kovarovic & Andrews, 2007). In fact, this has been previously observed in other studies, where different habitat categories overlapped (Degusta & Vrbra, 2005; Scott et al., 1999; Kovaric & Andrews, 2007; Curran & Haile-Selassie, 2016; Curran, 2012). Despite that, it was still possible to classify the habitat with a 71.2% percentage of correct classification on average.

Additionally, it was also apparent from the Decision Boundary plot that the assignment of the 'wet and dry' category falls towards habitat type. Naturally, soil moisture is deeply linked to vegetation type (Bond, 2008; Good & Caylor, 2011). Nonetheless, Kanapoi revealed slightly wetter environments than Allia Bay despite, in contrast, showing a slightly higher representation in open environments (open and light cover) than Allia Bay. In this regard, it might be important to include a more specific category about soil moisture in addition to habitat type, although the two are mutually implied. The most relevant variables for this classification, i.e., the radius of the distal articular end (Distrad) and the Functional length (B), were already shown to have the strongest relationship with individual habitat preference (Barr, 2014).

These results are different from those obtained by Dumouchel and colleagues (2021), who concluded that Allia Bay and Kanapoi have a higher representation of open habitats (47% and 61%, respectively). Yet, their analysis also indicates highly heterogeneous environments, as the remaining individuals are distributed in other categories. Although the fossil samples are not entirely equal because two fossil individuals were removed in this study, the difference between these results comes, almost certainly, from using models with different accuracies.

In their ecomorphological analysis, for habitat classification, Dumouchel and colleagues (2021) obtained an average accuracy of 77.8% using a linear discriminant function analysis (LDA), which is better than the results obtained here (71.2%). However, based on the description of their methods, I suspect that their results are overfitted. In their study, it is not mentioned which process they used to assess the quality of adjustment of the data, however, most ecomorphological studies (e.g. Kappelman, 1988, 1991; Kappelman et al., 1997; Kovarovic & Andrews, 2007; Plummer & Bishop, 1994) use the re-substitution process, which overestimates the classification power of canonical variables (Kovarovic et al., 2011). Cross-validation is significantly better for estimating the discriminative power of the data (Curran, 2012) and it was the approach preferred throughout this dissertation.

Furthermore, the use of different machine learning algorithms showed that LDA, a method often used in ecomorphological analyses, was not the model with the best accuracy results. Hence, when compared to other ecomorphological analyses, the results of this study prove to be more accurate for both habitat classification using astragalar data (e.g. Barr, 2015; DeGusta & Vrba, 2003), as well as other post-cranial bones (e.g. Kappelman, 1988, 1991; DeGusta & Vrba, 2005; Kovarovic & Andrews, 2007; Curran, 2012). As it was the case when analysing the ‘wet and dry’ categories, the results presented here are better, with a difference of only 1.5% as compared to the results obtained by Curran (2012) in their analysis of the third phalanx plantar margin in cervids (although the third phalanx has direct contact with the substrate, unlike the astragalus).

In most LDA applications, researchers tacitly assume that their data conform to the assumptions of the normal theory (Feldesman, 2002). However, only when the results are compared to another technique that requires fewer or different assumptions, it can be decided which technique fits the data better (Feldesman, 2002). The use and choice of several different algorithms, depending on the specific problem, are important to successfully solve group analysis and classification problems using morphometric data (Püschel et al., 2018), improving the traditional way classification tasks have been performed in these areas (Püschel et al., 2018). Nevertheless, these results could be improved with larger sample size, as presented by Plummer et al. (2008, 2015), whose result values are higher than these. Larger samples will allow more robust comparisons between different types of ecomorphic analyses (Plummer et al., 2008).

Ultimately, the random forests model is the machine learning approach that shows consistently high levels of accuracy. Other approaches, such as SVM and KNN, also showed good accuracy values, however, the models turned out to be statistically different and for this very reason the RF model was the chosen one. The disadvantage that comes with this accuracy of RF is the lack of interpretability and clarity of the model, often described as "black box" machine learning (i.e., algorithms where the internal parameters are essentially uninterpretable by humans) (Hosseini et al., 2020). Hence, one should take special concern with supervised applications, such as validating all the steps involved in the model design to obtain an unbiased estimate for accuracy (Tarca et al., 2007).

Finally, the biggest challenge faced when applying ecomorphological analysis is understanding what the results mean in terms of interpretable ecological variables (Barr, 2017).

This system only considers four habitat categories, with the intermediate categories (light and heavy cover) grouping several habitat types together. Thus, given the complexity of environments, this general classification obscures ecological nuances that exist in the real world (Kovarovic et al., 2018). Furthermore, using two categories to describe the moisture present in the soil is underestimating a certainly more complex phenomenon.

Ecometrics, on the other hand, has a greater emphasis on the multidimensional spectrum of many environments (Kovarovic et al., 2018). Compared to ecomorphological analysis, ecometric results provide a clearer assignment to a potential climate.

Ecometric results

Primarily, to build a good ecometric model, it is important to pay attention to the overall richness and the distribution pattern of the taxa that form the faunal community, as these are factors that affect the accuracy of predicting the correct habitat association for the fauna (Mares & Willig 1994; Andrews, 1996). Since 46 of the 55 species of bovids occurring in sub-Saharan Africa were used (Barr, 2017), this should not be a concern in the present study.

Here again, among the ML regression models included in this study, it was shown that the RF model is also the machine learning approach that shows consistently higher levels of accuracy when predicting the bioclimatic variables used in the test dataset. As with the ecomorphological analysis, it was also tested different machine learning algorithms, including the Linear model, although some of this model's assumption were violated, as the data is not independent (i.e., this sort of data is autocorrelated) (Serber & Lee, 2003). Yet, it was included for comparison to show the importance of testing more models beyond those commonly known and frequently used. Similarly, for the land cover classification, the RF model also shows the best results.

Moreover, it was also found that the variables with the most importance in the ecomorphological analysis are again the most relevant in the analysis of the functional morphology of the astragalus at the community level, as previously observed by Barr (2017). However, it should be noted that for some of the temperature variables, the Width at the level of flange on lateral surface (WAF) was also important, in addition to DistRad and B.

The results suggest that ecological information from mammal fauna can be successfully used to predict several African bioclimatic variables in addition to precipitation and land cover, as has been demonstrated previously in astragalar measurements in modern Africa (Barr, 2017). Among the eight bioclimatic variables included in this study, mean annual precipitation (R^2=0.88) and temperature seasonality (R^2=0.87) were the most accurately predicted by the regression models. On the other hand, the maximum temperature of the warmest month (R^2=0.70) and mean annual temperature (R^2=0.73) were the ones with the lower results, which means they do not reconstruct so accurately as the other variables, but the values are not substantially different. RMSE or MAE error metrics are not used in the comparison between different variables as they are not dimensionless because their value depends on the units in which the variable (i.e. precipitation or temperature) is measured.

Compared to the results from other ecometric studies for the bioclimatic variables, the results of this analysis are shown to be consistently better than previous works using other morphological features, such as dental traits (e.g., Žliobaitė et al., 2016; Faith et al., 2019; Liu et al., 2012) or locomotor aspects, such as tail length (e.g., Lawing et al., 2012) and calcaneal gear ratio (e.g., Polly, 2010). As for the land cover, as compared to other ecometric studies carried out using locomotor traits testing the same habitat categories as those tested in this analysis (e.g., Polly, 2010; Lawing et al., 2012; Barr, 2017; Short & Lawing, 2021), the results of this study are more accurate. The significance of the ecometric results presented here lies not only in the quantitative formalization of the relationship between the astragalus calculated at the community level and all environmental parameters tested but also in the application of different algorithms, which allow researchers to choose the most accurate one depending on the empirical results obtained from each specific study.

The models in this study show high overall success prediction and classification rates, which were also evident in the spatial prediction plots. The maps made using the astragalar variables of the comparative data were practically equal to the maps with the actual values. Meaning the RF model made a good association between the mean of each ecometric variable with the values of the bioclimatic variables and the land cover categories, in each grid cell on the map. Therefore, based on the metrics obtained, these results provide confidence in the values of the tested bioclimatic variables and the predicted habitat preference assignments for the Kanapoi and Allia Bay fossil samples.

According to the ecometric estimates, the two palaeontological sites follow relatively similar climatic patterns. The results show high mean temperatures between 23.8°C for Allia Bay and 25.7°C for Kanapoi, with minimum and maximum values within a range included in the tropical domain (Bonnefille, 2004). Furthermore, the value of temperature variation was low, which means that there was fewer extreme differences throughout the year, resulting in seasons not being pronounced, but rather a range of uniform and mild climate, as expected for a tropical climate region. It is commonly inferred that Africa was warmer in the past (Bobe et al., 2002; Fernández & Vrba, 2006; Potts, 2007; Passey et al., 2010; Fortelius et al., 2016). The last two studies (Passey et al., 2010; Fortelius et al., 2016), applied to the Turkana basin, also show this trend of absence of temperature variations.

On the other hand, precipitation values indicate a considerable difference between these two sites, with Kanapoi showing an average precipitation value of 1702.9 mm compared to 853.2 mm for Allia Bay. Within these obtained values there is a clear difference between the wettest and driest extremes for the two palaeontological sites, which is ultimately reflected in the precipitation seasonality values. According to Hare (2003), CV is used to classify the degree of variability of rainfall events as minor (CV <20), moderate (20 <CV <30) and high (CV> 30). Since the values were 58.2% for Allia Bay and 60.1% for Kanapoi, there was a change in the amount and/or pattern of rainfall during the time represented at the site by the fossils used. In the ecomophological analysis, a similar pattern occurred, with Kanapoi being wetter than Allia Bay, but the differences were minimal and both sites showed a mixture of the 'wet and dry' categories.

In general, most of Africa is characterised by a strong seasonal cycle in the precipitation regime and a pronounced dry season (Nicholson, 2000; Kingston, 2007). These rainfall patterns, especially in Tropical Africa, are mainly controlled by the strength of the Afro-Asian monsoon circulation and the migration of the Intertropical Convergence Zone (ITCZ)[3] (Nicholson, 2000; Hardt et al., 2007; Kingston, 2007; Geen et al., 2020). Thus, the timing of the monsoon rainfall increase in East Africa is primarily a product of the monsoons as localised seasonal migrations of the tropical convergence zone (Geen et al., 2020). The ITCZ follows the seasonal and latitudinal movement of incoming solar insolation (Lepre et al., 2007), changing

---

[3] ITCZ is defined as the location where the trade winds from the northern and southern hemispheres converge (Geen et al., 2020). The rising air in the ITCZ is cooled by expansion, resulting in condensation and rain. This zone of maximum rainfall follows the latitudinal position of the sun overhead, resulting in a general bimodal seasonal distribution of rainfall associated with the passage of the ITCZ (Kingston, 2007).

its mean position based on the temperature contrast between the northern and southern hemispheres, migrating towards the warming hemisphere (Schneider et al., 2014).

Therefore, these climate systems can change according to periodic variations in the Earth's orbit around the Sun, as these variations control the amount of incident solar radiation (insolation) (Zhang et al., 2015). These changes are the result of variations in orbital eccentricity, obliquity and, more specifically, the precession of the Earth's orbital axis (Prell & Kutzbach, 1987; Clement et al., 2004; Zhang et al., 2015). Precession is the orbital cycle (20,000 years) that determines the position of the seasons within the Earth's revolution (Berger et al., 1992), influenced by changes in Milankovitch periodicities (Berger et al., 1993).

Throughout the extent of the rift, without this climate control, East Africa cannot receive sufficient rainfall to fill large deep freshwater lakes (Maslin & Christensen, 2007). In fact, there is now considerable evidence for this precessional forcing on moisture availability in the tropics in East Africa during the Pliocene (Brown & Feibel, 1986; Feibel et al., 1989; Brown, 1995; deMenocal, 1995, 2004; Deino et al., 2006; Hopley et al., 2007; Kingston et al., 2007; Lepre et al., 2007). This oscillation in monsoon rainfall led to fluctuations within lakes (Deino et al. 2006, Kingston et al. 2007, Magill et al. 2013, Scholz et al. 2007, Wilson et al. 2014), changes in lake distribution (Trauth et al. 2007) and intervals of fluvial deposition (McDougall et al. 2005).

The Turkana Basin is linked to the large and ancestral Omo River (Brown & Feibel, 1988; Feibel, 1993; 2011). The amount of Omo River water that entered and transited the Turkana Basin for at least five million years (Levin et al., 2011) depended on rainfall in the Ethiopian highlands and the diversion of streams from the Omo River basin to the Nile (Brown & Feibel, 1991). In addition, the Kanapoi Formation lies in the floodplains of the lower Kerio River valley, which feed into the Omo-Turkana Basin (Dumouchel et al., 2021). Therefore, a large river such as the Omo can experience seasonal flooding altering habitats (Ward et al., 1999) and the corresponding lake level transgressions and regressions consisted of wetter and more seasonal conditions (Trauth et al., 2005, 2007).

This is supported by previous analyses of palaeoprecipitation for the Turkana Basin. For example, Fernandez and Vrba (2006) inferred a palaeoprecipitation of 1100 mm/year for 3.5 Ma, based on taxonomic analysis of macro mammal faunas. Similarly, the ecometric study by Fortelius et al. (2016) also confirmed these trends, according to their estimates, between 4-3

Ma, the basin had annual precipitation values ranging between ~ 600 and 1400mm. Moreover, Feibel (1988) had already found, based on palaeosol types, that by ~ 3.5 Ma the Turkana Basin was characterised by local heavy rainfall conditions. Bonnefille et al. (2004) inferred precipitation values between ~ 800 and 1200 mm / year, for the Pliocene. Further support for seasonality comes from mammal community structures (Reed, 1997; Bobe & Eck, 2001; Reed & Fish, 2005), palaeodiet reconstructions of hominins (Teaford & Ungar, 2000), vetissols (Wynn, 2000) and oxygen isotope ratios in fossil enamel of faunal teeth (Beasley et al., 2016).

Meanwhile, the higher mean precipitation signal for Kanapoi inferred in this study may be a consequence of the expansion of a large lake, as the presence of surface water appears to inflate local rainfall estimates significantly (Fortelius et al., 2016). Around 4.14 million years ago, the largest known lake in the history of the Omo-Turkana Basin, the Lonyumun palaeo-lagoon, emerged, extending from Kanapoi in the south to the Mursi region in the north (Feibel, 2011). The expansion of this lagoon undoubtedly had an important effect on the geography, vegetation, and fauna of the region (Bobe et al., 2020). Furthermore, the Kanapoi sedimentary sequence contains fluvial and deltaic deposits from before and after the Lake Lonyumun interval (Feibel et al., 1991; McDougall and Brown, 2008), deposited by the Kerio River (Feibel, 2003). As such, the palaeohydrology of the Turkana Basin over time was characterised by an alternating lake and riparian conditions (Feibel, 2011), ensuring permanent water abundance throughout climatic cycles (Manthi et al., 2020b; Joordens et al., 2019; Bobe et al., 2020). Nevertheless, these values represent an ‘average’ for the temporal frame represented by the fossils used and may not fully reflect Kanapoi's environmental variability.

It is essential to recognize that nature is highly complex and variable, particularly when situated in a context such as the Rift Valley. In addition to precipitation variation by the precession forcing, other processes may also have contributed to this ecological variability, including abiotic agents (topography, localized climatic conditions, volcanic disturbances, fires) as well as biotic factors (organisms and their resources mediation) (Kingston, 2007).

The floodplains, lake margins and volcanic terrains that prevail in the Rift Valley settings are unstable by nature (Levin, 2015). The geographical distribution of faults, tectonic movements and the resulting basin geometry and structure are important controls on the sedimentology and stratigraphy of the lake of East African settings (Tiercelin, 1990; Cohen et al., 1997; Scholz et al., 1998). In particular, the Turkana Basin is located at the intersection of two rift systems (Bosworth & Morley, 1994) and was affected by occasional tectonic activity, which disrupted outflow and resulted in short-lived temporary lakes (Brown & Feibel, 1991).

Consequently, the climate patterns are markedly complex and highly variable (Nicholson, 2000). And the consequences of these seasonal and humid periods are reflected in the vegetation type, as indicated in the land cover results. As with the ecomorphological analysis, the land cover categories reveal that no single category is predominant, but rather there is a variety of environments with relatively similar portions, emphasising heterogeneity. Both sites reveal a combination of habitats such as woody savannas, evergreen broadleaf forest and grasslands, only in different portions. Allia Bay has a high presence of open environments such as grasslands (38%), but also highly values for environments dominated by forest cover such as evergreen broadleaf forest (26%) and woody savannahs (17.2%), with the remaining categories showing values below 7.8%. On the other hand, Kanapoi shows more distributed values, with a slightly higher representation in environments with a higher presence of forest cover, such as woody savannas (27.4%) and evergreen broadleaf forest (22.2%), followed by more open environments, such as grasslands (17.8%), permanent wetlands (13.2%) and open shrublands (12.6%), among others (<3.4%). This means that the habitat of these sites was dominated by forest cover environments, which could exceed 60%, with heights greater than 2m, with trees and shrubs that remain green throughout the year (evergreen broadleaf forest) but that was also characterised by the presence of land with trees and herbaceous vegetation and other understory systems (woody savannahs) and more open systems, without so much forest cover (grasslands). For Kanapoi there is also some representation of bovids associated with permanent wetlands and open shrublands, characterised by land with a permanent mixture of water and land with shrub cover.

This combination of different vegetation types is expected for complex sites with multiple processes and agents capable of producing ecological heterogeneity (Kingston, 2007). Currently, there are some examples of these types of environments such as the miombo and mopane forests (Siyum, 2020) and the sub-Saharan savannahs of Africa (Prentice et al., 1992), with vegetation types of seasonal or semi-dry tropical forests, where the dominant trees are green with tree cover from continuous (in true forests) to sparse (in some savannas) (Prentice et al., 1992). These types of environments, also known as seasonally dry tropical forests, are defined as a forest with 500-2000 mm of annual precipitation and a pronounced dry season

(Miles et al., 2006). After all, each biome (e.g. forest, savannah, grassland) is distributed over a wide range of precipitation (Staver et al., 2011; Hirota et al., 2011; Guan et al., 2014).

These results are congruent with palaeoenvironmental reconstructions previously made. The literature of Kanapoi's environmental reconstruction has indicated open and closed, with wooded and grassy environments (Harris et al., 2003; Feibel et al., 1991; Feibel, 2003; Wynn, 2000; Quinn & Lepre, 2020; Bobe et al., 2020), with a fraction of woody canopy cover between 40% and 60% (Cerling et al., 2011; Manthi et al., 2020a, 2020b). Pedogenic carbonate isotope studies reveal high isotopic variability over relatively short distances (Wynn, 2004; Quinn & Lepre, 2020), consistent with C3 and C4 mosaic landscapes. Recent studies at Kanapoi have shown the presence of a high diversity of very large mammals (Bobe et al., 2020; Sanders, 2020), which would have a significant impact on the opening forest landscape (Bobe et al., 2020), contributing to palaeoenvironmental heterogeneity. In addition, stable isotopes of macro-and micro-mammal dentitions (Manthi et al., 2020a) also point to an irregular habitat.

Previous reconstructions in Allia Bay follow the same scenario, indicating a mixture of habitats, including habitat from closed gallery forest environments, but also floodplain grassland and marginal dry forest (Feibel et al., 1991; Coffing et al., 1994; Leakey et al., 1995; Schoeninger & Reeser, 1999), with seasonal environments (Macho et al., 2003), but with a greater emphasis on forest components (Schoeninger et al., 2003).

Studies for the overall Pliocene also indicate the presence of C3 vegetation by stable isotopes of herbivore enamel before 2.9 Ma (Bibi et al., 2013; Negash et al., 2015), as well as by soil carbonate records from vegetation in East and Central Africa (Cerling et al., 1977). In addition, flora characteristic of the modern West African rainforest is found in East Africa around 3.4 Ma (Bonnefille & Letouzey, 1976; Bonnefille, 1987). Early Pliocene sites suggest the presence of woodland/closed forests to moderate forests and woodland/grasslands (WoldeGabriel et al., 1994).

Whilst the results are in line with what has been previously stated, it is important to consider that the assemblage of fossil specimens analysed for Kanapoi, and especially Allia Bay is very limited and may not include a representative sample of these two sites. For example, in the study of Dumouchel and colleagues (2021) is mentioned that the Kanapoi sample includes several small-bodied taxa. Therefore, it could be that bovids from the forest are overrepresented. Nonetheless, poorly sampled fossil assemblages still provide significant ecometric patterns

(Vermillion et al., 2018), but this does not necessarily imply that sampling is not important (Faith et al., 2019).

Combined ecological interpretations of the Kanapoi and Allia Bay sites suggest that early hominids were drawn to a mosaic of environments (Wynn, 2000). It seems clear that *Australopithecus anamensis* thrived in these types of environments. *A. anamensis* lived in mosaic forests in the Afar Basin (White et al., 2006; Kullmer et al., 2008; Curran & Haile-Selassie, 2016; Saylor et al., 2019; Quinn & Lepre, 2020). Analyses of diet in Bovid tribes in Aramis (White et al., 2009b), Woranso-Mille (Curran & Haile-Selassie, 2016) and the Mursi Formation (Drapeau et al., 2014) exhibit a more browse-dominated signal. The Allia Bay and Kanapoi Bovidae datasets include more individuals with mixed feeding within each tribe (Dumouchel et al., 2021). However, in Kanapoi, primates, micro-mammals and hominins have emphasized mainly C3 resources, such as fruits and tree leaves (Sponheimer et al., 2006a, 2006b; Stewart & Rufolo, 2020), despite the availability of C4 resources exploited by macro- and micro-mammalian herbivores and *Theropithecus* (Manthi et al., 2020a). Hence, it seems that *A. anamensis* maintained adaptations to an arboreal lifestyle compared to most later hominids (Green et al., 2007; Ward et al., 2013). Furthermore, it is now considered that bipedalism probably arose in a woodier context (e.g., Prost, 1980; Moore, 1992; Beasley & Schoeninger, 2015; Senut et al., 2018; Kimura, 2019).

In general, the palaeoecological reconstructions of this study extend what is currently known of the habitats in which *A. anamensis* evolved, a heterogeneous but relatively open forest environment, with more or fewer patches of grassy areas shaped by changing hydrological conditions as well as the abundance of megaherbivores. The ability of the first *Australopithecus* to thrive in these dynamic environments seems to be an early indication of the adaptability and versatility of this genus (Manthi et al., 2020b).

It is now known that the environmental niche of hominids cannot be fixed to a particular vegetation type (Chase, 2011). In recent years, there has been a re-evaluation of the relationship between environment and human evolution (Maslin & Christensen, 2007). For example, it is no longer taken for granted that savannahs were the predominant environment of early hominin evolution (e.g., Vrba, 1985; Reed, 1997; Lee-Thorp et al., 2007), or that human evolution was forced by rapid increases in Africa's aridity (e.g., Vrba's Pulse Rotation Hypothesis, 1985, 1995b). Increased knowledge of African paleoclimates has brought new scenarios of human evolution, such as the Variability Selection Hypothesis (Potts, 1996, 1998a), which emphasizes the importance of climate instability as a mechanism for selection.

Potts (1998a, 1998b) suggest that the complex intersection of orbitally forced changes in insolation and the Earth's intrinsic feedback mechanisms result in inconsistent and extreme environmental variability to behavioural and morphological mechanisms, providing a significant evolutionary driver (deMenocal, 2004; Kingston, 2007). Several of the major events in the evolution of African fauna occurred when there were increases in the amplitudes of palaeoclimatic variability (Potts, 1998a). For example, global climate change around 2.8 million years ago (Ma), with the onset and intensification of high-latitude glacial cycles that overlapped with changes in African monsoon precipitation due to orbital insolation (deMenocal, 2004), has been implicated in episodes of late Pliocene faunal renewal that included the origin of the genus *Homo* (deMenocal, 2004, 2011).

Several lines of evidence from hominid sites in East Africa support this hypothesis, e.g., pollen sequence studies (Bonnefille, 1995), isotopic data from palaeosol carbonates (Cerling, 1992; Levin et al., 2004; Wynn, 2004) and renewal rates for large mammal taxa (Bobe et al., 2002; Bobe & Behrensmeyer, 2004; Fernández & Vrba, 2006). Habitat instability is increasingly considered an influence on human evolution and part of becoming human might involve adaptations to unpredictable and unstable environments (Kingston, 2007; Antón et al., 2014; Potts, 2013).

Although the results obtained here are highly informative, maximum confidence comes from combining analyses of whole faunal and floral assemblages with geological, geochemical, and other types of contextual data (Croft et al., 2018). Therefore, an ecometric analysis can certainly provide a very useful tool to complement and contribute to applied palaeoenvironment reconstructions in Africa. Taxon-free methods have advantages since they do not require the destruction of fossils specimens, such as when using chemical approaches (e.g., stable isotope analyses) (Žliobaitė, 2018) and can be used on fossils for which a taxonomic identification is not possible. Besides, with the advance of computer techniques and the development of packages such as *caret,* the use of these methods has become easier and more user-friendly.

Finally, it is also worth remembering that climate was only one of many factors affecting human evolution. Biological processes including genetic innovation, competition between species and dispersal capacity may also have played defining roles (Potts, 2007). As well as cultural factors that have increasing importance in later stages of human evolution (Boyd & Richerson, 1985; Boyd, 2018; Bender, 2020).

5) Conclusion

This study showed that taxon-free methods offer a good alternative when analysing commonly preserved and well represented fauna at sites such as Allia Bay and Kanapoi, as they provide high success rates of classification and prediction of the environments of interest. By using different ML algorithms, it was possible to produce robust models with high accuracy, which became clear when the results were compared with other studies using more traditional methods. Still, it is likely that more advanced techniques to capture morphology, such as GM, will improve the prediction of the models.

Based on prediction accuracy and visual assessment, ecometric methods seem more promising for palaeoenvironmental reconstruction, as compared to more traditional ecomorphological analyses. Nevertheless, it is still possible through the individual functional characteristics to obtain a complex relationship with the environment, as seen with the fact that the results between the two analyses are corresponding. Furthermore, the functional morphology of astragalus was shown to be consistent across all bioclimatic variables tested ($R^2 > 0.70$), in addition to land cover and mean annual precipitation, both previously tested.

The results are consistent with former reconstructions. Kanapoi and Allia Bay, during the period to which the fossils analysed relate, were warm, humid, but seasonal, containing a spectrum of different habitats, with a representation of forest as well as more open habitats. This seems to support the suggestion that *Australopithecus anamensis* thrived in highly heterogeneous environments that were constantly fragmented in response to environmental fluctuations related to orbital forcing, local tectonics and global climatic events/trends.

Future work on ecomorphology should seek to infer climate and biological change using ecometrics at the community level. I expect this work to contribute to this new direction of ecometrics research, possibly repeating analysis incorporating additional mammal fossils to improve the accuracy and precision of environmental reconstruction of Turkana Basin, an area of interest to a broad paleoanthropological community due to its great significance for human evolution. Finally, it would be interesting to extend ecometric analyses across different geographic and temporal scales associated with the history of our own evolution in order to contribute towards a better understanding of how climate has affected our evolutionary history, as well as providing additional future insights to understand climate change.

References

Agustí, J. (2007). The Biotic Environments of the Late Miocene Hominids. In Henke, W., Hardt, T., Tattersall, I. (Eds.). *Handbook of Paleoanthropology Volume II, Primate Evolution and Human Origins* (pp. 979-1010). Berlin, Germany: Springer-Verlag.

Anderson, J. (2013). Tragelaphus angasii Nyala. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.148-152). London, United Kingdom: Bloomsbury Publishing.

Andrews, P. (1996). Paleoecology and hominoid palaeoenvironments. *Biological Reviews, 71*, 257–300. DOI: 10.1111/j.1469-185X.1996.tb00749.x

Andrews, P. (2018). Foreword. In Croft, D.A., Su, D.F., Simpson, S.W. (eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp. V-VIII). Cham, Switzerland: Springer.

Andrews, P., Hixson, S. (2014). Taxon-free methods of palaeoecology. *Annales Zoologici Fennici, 51*, 269–284. DOI: 10.5735/086.051.0225.

Antón, S.C., Potts, R., Aiello, L.C. (2014). Evolution of early Homo: an integrated biological perspective. *Science, 345 (Issue 6192)*, 1236828. DOI: 10.1126/science.1236828.

Bailey, I.W., Sinnott, E.W. (1915). A botanical index of Cretaceous and Tertiary climates. *Science, 41*, 831–834. DOI: 10.1126/science.41.1066.831.

Bamford, M., Senut, B., Pickford, M. (2013). Fossil leaves from Lukeino, a 6-million-year-old Formation in the Baringo Basin, Kenya. *Geobios, 46*, 253–272. DOI: 10.1016/j.geobios.2013.02.001.

Barr, W.A. (2014). Functional morphology of the bovid astragalus in relation to habitat: controlling phylogenetic signal in ecomorphology. *Journal of Morphology, 275,* 1201–1216. DOI: 0.1002/jmor.20279

Barr, W.A. (2015). Paleoenvironments of the Shungura Formation (Plio-Pleistocene: Ethiopia) based on ecomorphology of the bovid astragalus. *Journal of Human Evolution*, *88,* 97–107. DOI: 10.1016/j.jhevol.2015.05.002

Barr, W.A. (2017). Bovid locomotor functional trait distributions reflect land cover and annual precipitation in sub-Saharan Africa. *Evolutionary Ecology Research, 18*, 253–269.

Barr, W.A. (2018). Ecomorphology. In Croft, D., Simpson, S., Su, D. (Eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp. 339-349). Cham, Switzerland: Springer.

Barr, W.A., Biernat, M. (2020). Mammal functional diversity and habitat heterogeneity: Implications for hominin habitat reconstruction. *Journal of Human Evolution, 146,* 102853. DOI: 10.1016/j.jhevol.2020.102853.

Barthelme, S. (2021). imager: Image Processing Library Based on 'CImg'. R package version 0.42.8. DOI: https://CRAN.R-project.org/package=imager

Beasley, M.M., Schoeninger, M.J. (2015). Miombo woodlands and early hominins: A comparison of carbonate stable isotope data from modern Koobi Fora and 3.97Ma Allia Bay fauna. *American Journal of Physical Anthropology, 156(60),* 81-81.

Beasley, M.M., Orland, I.J., Valley, J.W., Schoeninger, M.J. (2016). Seasonal variation in rainfall at Allia Bay, Kenya 3.97 Ma. *American Journal of Physical Anthropology, 159 (62),* 88-88.

Behrensmeyer, A.K., (1975). The taphonomy and paleoecology of Plio-Pleistocene vertebrate assemblages east of Lake Rudolf, Kenya. *Bulletin of The Museum of Comparative Zoology, 146*, 473-578. DOI: 10.5962/BHL.PART.22969

Behrensmeyer, A.K. (2006). Climate change and human evolution. *Science, 311*, 476–478. DOI: 10.1126/science.1116051.

Behrensmeyer, A.K., Reed, K.E. (2013). Reconstructing the habitats of *Australopithecus*: paleoenvironments, site taphonomy, and faunas. In: Reed, K., Fleagle, J., Leakey, R. (Eds.), *The Paleobiology of Australopithecus* (pp. 41-60). Dordrecht, Netherlands: Springer.

Behrensmeyer, A.K., Todd, N.E., Potts, R., McBrinn, G.E. (1997). Late Pliocene faunal turnover in the Turkana Basin, Kenya and Ethiopia. *Science, 278*, 1589-1594. DOI: 10.1126/science.278.5343.1589.

Behrensmeyer, A.K., Bobe, R., Alemseged, Z. (2007). Approaches to the analysis of faunal change during the East African Pliocene. In Bobe, R., Alemseged, Z., Behrensmeyer, A.K. (Eds.). *Hominin environments in the East African Pliocene: an assessment of the faunal evidence* (pp. 1-24). Dordrecht, Netherlands: Springer.

Ben-Hur, A., Ong, C.S., Sonnenburg, S., Schölkopf, B., Rätsch, G. (2008). Support vector machines and kernels for computational biology. *PLoS Computational Biology, 4,* e1000173. DOI: 10.1371/journal.pcbi.1000173.

Benatti, N. (2019), A machine learning approach to outlier detection and imputation of missing data. In Fisher, I. (Ed.). *Are post-crisis statistical initiatives completed?* (vol 49). Bank for International Settlements. DOI: https://EconPapers.repec.org/RePEc:bis:bisifb:49.

Bender, A. (2020). The Role of Culture and Evolution for Human Cognition. *Topics in Cognitive Science.* doi:10.1111/tops.12449

Benítez, H., Püschel, T.A. (2014). Modelando la varianza de la forma: morfometría geométrica aplicaciones en biología evolutiva. *International Journal of Morphology, 32(3)*, 998-1008. DOI: 10.4067/S0717-95022014000300041

Berger, A., Loutre, M.F., Laskar, J. (1992). Stability of the astronomical frequencies over the Earth's history for paleoclimate studies. *Science, 255*, 560-566. DOI: 10.1126/science.255.5044.560

Berger, A., Loutre, M.F., Tricot, C. (1993). Insolation and Earth's orbital periods*.* *Journal of Geophysical Research, 98*, 10341–10362. DOI: 10.1029/93JD00222

Berger, L.R., de Ruiter, D.J., Churchill, S.R., Schmid, P., Carlson, K.J., Dirks, P.H., Kibii, J.M. (2010). *Australopithecus sediba*: a new species of Homo-like australopith from South Africa*.* *Science, 328*, 195–204. DOI: 10.1126/science.1184944.

Bibi, F., Souron, A., Bocherens, H., Uno, K., Boisserie, J.R. (2013). Ecological change in the lower Omo Valley around 2.8 Ma. *Biology Letters, 9*, 20120890. DOI: 10.1098/rsbl.2012.0890

Bishop, L., King, T., Hill, A., Wood, B. (2006). Palaeoecology of Kolpochoerus heseloni (= K. limnetes): a multiproxy approach. *Transactions of the Royal Society of South Africa, 61*, 81-88. DOI: 10.1080/00359190609519956.

Bivand, R., Lewin-Koh, N. (2021). maptools: Tools for Handling Spatial Objects. R package version 1.1-1. DOI: https://CRAN.R-project.org/package=maptools

Bivand, R., Rundel, C. (2020). rgeos: Interface to Geometry Engine – Open Source ('GEOS'). R package version 0.5-5. DOI: https://CRAN.R-project.org/package=rgeos

Bivand, R., Keitt, T., Rowlingson, B. (2021). rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-23. DOI: https://CRAN.R-project.org/package=rgdal

Blumenthal, S.A., Levin, N.E., Brown, F.H., Brugal, J., Chritz, K.L., Harris, J.M., Jehle, G.E., Cerling, T.E. (2017). Aridity and hominin environments. *PNAS*, *114* (28), 7331–7336. DOI: 10.1073/pnas.1700597114.

Bobe, R. (2006). The evolution of arid ecosystems in Eastern Africa. *Journal of Arid Environments, 66*, 564-584. DOI: 10.1016/j.jaridenv.2006.01.010.

Bobe, R. (2011) Fossil mammals and paleoenvironments in the Omo-Turkana Basin. *Evolutionary Anthropology, 20,* 254 - 263. DOI:10.1002/evan.20330.

Bobe, R., Eck, G.G. (2001). Response of African bovids to Pliocene climatic change. *Paleobiology, 27*, 1-47. DOI: 10.1666/0094-8373(2001)027<0001:ROABTP>2.0.CO;2

Bobe, R., Behrensmeyer, A.K. (2004). The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus Homo. *Palaeogeography, Palaeoclimatology, Palaeoecology, 207*, 399-420. DOI: 10.1016/j.palaeo.2003.09.033.

Bobe, R., Leakey, M.G. (2009). Ecology of Plio-Pleistocene mammals in the Omo-Turkana Basin and the emergence of Homo. In Grine, F.E., Fleagle, J.G., Leakey, R.E. (eds.). *The first humans: origin and early evolution of the genus Homo* (pp. 173-184). Berlin, Germany: Springer.

Bobe, R., Behrensmeyer, A.K., Chapman, R.E. (2002). Faunal change, environmental variability and late Pliocene hominin evolution. *Journal of Human Evolution, 42*, 475–497. DOI: 10.1006/jhev.2001.0535.

Bobe, R., Alemseged, Z., Behrensmeyer, A.K. (Eds.). (2007a). *Hominin Environments in the East African Pliocene: an Assessment of the Faunal Evidence.* Dordrecht, Netherlands: Springer.

Bobe, R., Behrensmeyer, A.K., Eck, G.G., Harris, J.M. (2007b). Patterns of abundance and diversity in late Cenozoic bovids from the Turkana and Hadar basins, Kenya and Ethiopia. In Bobe, R., Alemseged, Z., Behrensmeyer, A.K. (Eds.). *Hominin environments in the East African Pliocene: an assessment of the faunal evidence* (pp. 129-158). Dordrecht, Netherlands: Springer.

Bobe, R., Manthi, F., Ward, C.V., Plavcan, J.M., & Carvalho, S. (2020). The ecology of *Australopithecus anamensis* in the early Pliocene of Kanapoi, Kenya. *Journal of human evolution, 140*, 102717. DOI: 10.1016/j.jhevol.2019.102717.

Bock, W.J. (1989). From biologische Anatomie to ecomorphology. *Netherlands Journal of Zoology, 40,* 254–277. DOI: 10.1163/156854289X00291

Böhme, M., Nikolay, S., Fuss, J., Tröscher, A., Deane, A., Prieto, J., Kirscher, U., Lechner, T., Begun, D. (2019). A new Miocene ape and locomotion in the ancestor of great apes and humans. *Nature, 575*, 1-5. DOI: 10.1038/s41586-019-1731-0.

Bond, W.J. (2008). What limits trees in C4 grasslands and savannas? *Annu. Rev. Ecol. Evol. Syst., 39*, 641–659. DOI: 10.1146/annurev.ecolsys.39.110707.173411

Bonnefille, R. (1995). A reassessment of the Plio-Pleistocene pollen record of East Africa. In Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (eds.). *Paleoclimate and evolution with emphasis on human origins* (pp. 299–310). New Haven, CT: Yale University Press.

Bonnefille, R., Letouzey, R. (1976). Fruits fossiles d'*Antrocaryon* dans la vallée de l'Omo (Ethiopie). *Adansonia, 16*, 65–82.

Bonnefille, R., Potts, R., Chalié, F., Jolly, D., Peyron, O. (2004). High-resolution vegetation and climate change associated with Pliocene *Australopithecus afarensis*. *PNAS, 101(33)*, 12125-12129. DOI: 10.1073/pnas.0401709101.

Bookstein, F.L. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology.* Cambridge, UK: University Press.

Bosworth, W., Morley, C.K. (1994). Structural and stratigraphic evolution of the Anza Rift, Kenya. *Tectonophysics, 236*, 93-115. DOI: 10.1016/0040-1951(94)90171-6

Boyd, R. (2018). Not by Brains Alone: The Vital Role of Culture in Human Adaptation. In Boyd, R. (Ed.). *A different kind of animal: how culture transformed our species* (pp. 9-62). Oxford, UK: Princeton University Press.

Boyd, R., Richerson, P.J. (1985*). Culture and the evolutionary process.* Chicago, IL: University of Chicago Press.

Boyer, D.M., Yapuncich, G.S., Butler, J.E., Dunn, R.H., Seiffert, E.R. (2015). Evolution of postural diversity in primates as reflected by the size and shape of the medial tibial facet of the talus. *American journal of physical anthropology, 157(1),* 134–177. DOI: 10.1002/ajpa.22702

Braun, D.R., Harris, J.W.K., Levin, N.E., McCoy, J.T., Herries, A.I.R., Bamford, M.K., Bishop, L.C., Richmond, B.G., Kibunjia, M. (2010). Early hominin diet included diverse terrestrial and aquatic animals 1.95 Ma in East Turkana, Kenya. *PNAS, 107 (22),* 10002-10007. DOI: 10.1073/pnas.100218110.

Breiman, L. (1984). *Classification and regression trees*. New York, NY: Routledge

Brown, F.H. (1995). The Turkana Basin and Paleoclimatic Reconstruction in East Africa. In Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (eds.). *Paleoclimate and evolution with emphasis on human origins* (pp.319–330). New Haven, CT: Yale University Press.

Brown, F.H., Feibel, C.S. (1986). Revision of the lithostratigraphic nomenclature in the Koobi Fora region, Kenya. *Journal of the Geological Society, 143*, 297-310. DOI:10.1144/gsjgs.143.2.0297

Brown FH, Feibel CS. (1988). “Robust” hominids and Plio-Pleistocene paleogeography of the Turkana Basin, Kenya and Ethiopia. In Grine, F.E. (ed.). *Evolutionary History of the “Robust” Australopithecines* (pp. 325–41). New York, NY: de Gruyter.

Brown, F.H., Feibel, C.S. (1991). Stratigraphy, depositional environments, and paleogeography of the Koobi Fora Formation. In Harris, J.M. (Ed.) *Koobi Fora Research Project, Vol. 3* (pp. 1–30). Oxford: Clarendon Press.

Broxton, P.D., Zeng, X., Sulla-Menashe, D., Troch, P.A. (2014). A global land cover climatology using MODIS data. *Journal of Applied Meteorology and Climatology*, *53*, 1593–1605. DOI: 10.1175/JAMC-D-13-0270.1

Brunet, M., Beauvilain, A., Coppens, Y., Heintz, E., Moutaye, A.H.E., Pilbeam, D. (1996). *Australopithecus bahrelghazali*, a new species of early hominid from Koro Toro region, Chad. *Comptes Rendus de l’Académie des Sciences, Série 2, Sciences de la Terre et des Planètes, 322*, 907–913.

Brunsdon, C., Chen, H. (2014). GISTools: Some further GIS capabilities for R. R package version 0.7-4. DOI: https://CRAN.R-project.org/package=GISTools

Burnaby, T. (1966). Growth-Invariant Discriminant Functions and Generalized Distances. *Biometrics, 22*(1), 96-110. DOI: 10.2307/2528217.

Campisano, C.J., Cohen, A.S., Arrowsmith, J.R., Asrat, A., Behrensmeyer, A.K., Brown, E.T., Deino, A., Deocampo, D., Feibel, C.S., Kingston, J.D., Lamb, H.F., Lowenstein, T.K., Noren, A., Olago, D.O., Owen, R.B., Pelletier, J.D., Potts, R., Reed, K.E., Renaut, R.W., Russell, J.M., Russell, J.L., Schäbitz, F., Stone, J.R., Trauth, M.H., Wynn, J.G. (2017). The Hominin Sites and Paleolakes Drilling Project: high-resolution paleoclimate records from the East African Rift System and their implications for understanding the environmental context of hominin evolution. *PaleoAnthropology, 2017*, 1-43. DOI: 10.4207/PA.2017.ART104

Cano, A.G, Fernández, M.H., Álvarez-Sierra, Á. (2013). Dietary Ecology of Murinae (Muridae, Rodentia): A Geometric Morphometric Approach. *PloS one, 8,* e79080. DOI: 10.1371/journal.pone.0079080.

Carrión, J.S., Scott, L., deMenocal, P. (2019). Paleofloras, paleovegetation and human evolution. *Review of Palaeobotany and Palynology*, *267,* 32–38. DOI: 10.1016/j.revpalbo.2019.04.006.

Caruana, R., Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd international conference on Machine learning, (pp. 161–168). ACM.

Cassinelo, J. (2013). Ammotragus lervia Aoudad (Barbary Sheep, Arui). In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.595-598). London, United Kingdom: Bloomsbury Publishing.

Castelló, J. (2016). *Bovids of the World Antelopes, Gazelles, Cattle, Goats, Sheep, and Relatives*. Princeton: Princeton University Press.

Cerling, T.E. (1992). Development of grasslands and savannas in East Africa during the Neogene. *Palaeogeography, Palaeoclimatology, Palaeoecology, 97,* 241–47. DOI: 10.1016/0031-0182(92)90211-M

Cerling, T.E., Hay, R.L., O'Neil, J.R. (1977). Isotopic evidence for dramatic climatic changes in East-Africa during Pleistocene. *Nature, 267,*137–138. DOI: 10.1038/267137a0

Cerling, T.E., Wynn, J.G., Andanje, S.A., Bird, M.I., Korir, D.K., Levin, N.E., Mace, W., Macharia, A.N., Quade, J., Remien, C.H. (2011). Woody cover and hominin environments in the past 6 million years. *Nature, 476*, 51–6. DOI: 10.1038/nature10306.

Cerling, T.E., Manthi, F.K., Mbua, E.N., Leakey, L.N., Leakey, M.G., Leakey, R.E., Brown Francis, H., Grine, F.E., Hart, J.A., Kalemeg, P., Roche, H., Uno, K.T., Wood, B.A. (2013). Stable isotope-based diet reconstructions of Turkana Basin hominins. *PNAS, 110,* 10501–10506. DOI: 10.1073/pnas.1222568110.

Cerling, T.E., Andanjec, S.A., Blumenthal, S.A., Brown, F.H., Chritz, K.L., Harris, J.M., Hart, J.A., Kirera, F.M., Kaleme, P., Leakey, L.N., Leakey, M.G., Levin, N.E., Manthi, F.K., Passey, B.H., Uno, K.T. (2015). Dietary changes of large herbivores in the Turkana basin. Kenya from 4 to 1 ma. *PNAS, 112*(37), 11467–11472. DOI: 0.1073/pnas.1513075112.

Chai, T., Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. *Geoscientific Model Development, 7,* 1247-1250. DOI: 10.5194/gmd-7-1247-2014.

Chase, J.M. (2011). Ecological niche theory. In: Scheiner, S.M., Willig, M.R. (Eds.), *The Theory of Ecology* (pp. 93-107). Chicago, IL: University of Chicago Press.

Chorowicz, J. (2005). The East African Rift System. *Journal of African Earth Sciences, 43,* 379-410. DOI: 10.1016/j.jafrearsci.2005.07.019.

Clement, A.C., Hall, A., Broccoli, A.J. (2004). The importance of precessional signals in the tropical climate. *Climate Dynamics, 22,* 327-341. DOI: 10.1007/s00382-003-0375-8

Coffing, K., Feibel, C., Leakey, M., Walker, A. (1994). Four-million-year-old hominids from East Lake Turkana, Kenya. *American journal of physical anthropology*, *93*(1), 55–65. DOI: 10.1002/ajpa.1330930104.

Cohen, A.S., Lezzar, K.E., Tiercelin, J.-J., Soreghan, M. (1997). New paleogeographic and lake-level reconstructions of Lake Tanganyika: implications for tectonic, climatic and biological evolution in a rift lake. *Basin Research, 9,* 107-132. DOI: 10.1046/j.1365-2117.1997.00038.x

Colangelo, P., Castiglia, R., Franchini, P., Solano, E. (2010). Pattern of shape variation in the eastern African gerbils of the genus Gerbilliscus (Rodentia, Muridae): Environmental correlations and implication for taxonomy and systematics. *Mammalian Biology, 75*, 302-310. DOI: 10.1016/j.mambio.2009.05.001.

Cooke, S. (2011). Paleodiet of Extinct Platyrrhines With Emphasis on the Caribbean Forms: Three-Dimensional Geometric Morphometrics of Mandibular Second Molars. *Anatomical record, 294,* 2073-91, DOI: 10.1002/ar.21502.

Croft, D., Simpson, S., Su, D. (2018). Introduction to Paleoecological Reconstruction. In Croft, D., Simpson, S., Su, D. (Eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp. 1-7). Cham, Switzerland: Springer.

Curran, S.C. (2012). Expanding ecomorphological methods: geometric morphometric analysis of Cervidae post-crania. *Journal of Archaeological Science, 39*, 1172–1182. DOI: 10.1016/j.jas.2011.12.028

Curran, S.C. (2015). Exploring Eucladoceros ecomorphology using geometric morphometrics. *Anatomical Record, 298,* 291–313. DOI: 10.1002/ar.23066.

Curran, S.C. (2018). Three-Dimensional Geometric Morphometrics in Paleoecology. In Croft, D., Simpson, S., Su, D. (Eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp.319-338). Cham, Switzerland: Springer.

Curran, S.C., Haile-Selassie, Y. (2016). Paleoecological reconstruction of hominin- bearing middle Pliocene localities at Woranso-Mille, Ethiopia. *Journal of Human Evolution, 96,* 97-112. DOI: 10.1016/j.jhevol.2016.04.002

Damuth, J.D., 1992. Taxon-free characterization of animal communities. In: Behrensmeyer, A.K., Damuth, J.D., DiMichele, W.A., Potts, R., Sues, J., Wing, S.L. (Eds.), Terrestrial Ecosystems Through Time: Evolutionary Paleoecology of Terrestrial Plants and Animals (pp. 183–203). Chicago, IL: University of Chicago Press.

Darwin, C. (1871). *The descent of man: And selection in relation to sex*. London, United Kingdom: John Murray.

David, J., Lloyd, P. (2013). *Damaliscus pygargus* Bontebok / Blesbok. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.496-501). London, United Kingdom: Bloomsbury Publishing.

Decoste, D., Schölkopf B. (2002). Training invariant support vector machines. *Mach. Learn, 46,* 161–190. DOI:10.1023/A:1012454411458

DeGusta, D., Vrba, E. (2003). A method for inferring paleohabitats from the functional morphology of bovid astragali. *Journal of Archaeological Science*, *30,* 1009–1022. DOI: 10.1016/S0305-4403(02)00286-8

Degusta, D., Vrba, E. (2005). Methods for inferring paleohabitats from the functional morphology of bovid phalanges. *Journal of Archaeological Science, 32*, 1099-1113. DOI: 10.1016/j.jas.2005.02.010

Deino, A.L., Kingston, J.D., Glen, J.M., Edgar, R.K., Hill, A. (2006). Precessional forcing of lacustrine sedimentation in the late Cenozoic Chemeron Basin, Central Kenya Rift, and calibration of the Gauss/Matuyama boundary. *Earth and Planetary Science Letters, 247*, 41-60. DOI: 10.1016/j.epsl.2006.04.009

deMenocal, P.B. (1995). Plio-Pleistocene African climate. *Science, 270,* 53-59. DOI: 10.1126/science.270.5233.53

deMenocal, P.B. (2004). African climate change and faunal evolution during the Pliocene- Pleistocene. *Earth and Planetary Science Letters, 220*, 3–24. DOI: 10.1016/S0012-821X(04)00003-2.

deMenocal, P.B. (2011). Climate and human evolution. *Science, 331*, 540-542. DOI: 10.1126/science.1190683.

Domínguez-Rodrigo M. (2014). Is the "Savanna Hypothesis" a dead concept for explaining the emergence of the earliest hominins*? Current Anthropology, 55*, 59–81. DOI: 10.1086/674530.

Drapeau, M.S., Bobe, R., Wynn, J.G., Campisano, C.J., Dumouchel, L., Geraads, D. (2014). The Omo Mursi Formation: A window into the East African Pliocene. *Journal of Human Evolution, 75*, 64-79. DOI: 10.1016/j.jhevol.2014.07.001

Du, A., Rowan, J., Wang, S.C., Wood, B.A., Alemseged, Z. (2020). Statistical estimates of hominin origination and extinction dates: a case study examining the *Australopithecus anamensis–afarensis* lineage. *Journal of human evolution, 138*, 102688. DOI: 10.1016/j.jhevol.2019.102688

Dumouchel, L., Bobe, R., Wynn, J. G., Barr, W.A. (2021). The environments of *Australopithecus anamensis* at Allia Bay, Kenya: A multiproxy analysis of early Pliocene Bovidae. *Journal of human evolution*, *151*, 102928. DOI: 10.1016/j.jhevol.2020.102928

Duncan, P. (2013). *Damaliscus lunatus* Topi / Tsessebe / Tiang / Korrigum. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.502-510). London, United Kingdom: Bloomsbury Publishing.

Elton, S. (2002). A reappraisal of the locomotion and habitat preference of Theropithecus oswaldi. *Folia Primatologica, 73*, 252–280. DOI: 10.1159/000067457.

Elton, S. (2008). The environmental context of human evolutionary history in Eurasia and Africa. *Journal of Anatomy, 212(4), 377–393.* DOI: 10.1111/j.1469-7580.2008.00872.x.

Elton, S., Jansson, A.-U., Meloro, C., Louys, J., Plummer, T.W., Bishop, L.C. (2016). Exploring morphological generality in the Old World monkey postcranium using an ecomorphological framework. *Journal of Anatomy, 228,* 534–560. DOI: 10.1111/joa.12428.

Eronen, J.T., Polly, P.D., Fred, M., Damuth, J., Frank, D.C., Mossbrugger, V., Scheidegger, C., Stenseth, N.C., Fortelius, M. (2010a). Ecometrics: the traits that bind the past and present together. *Integrative Zoology, 5*, 88–101. DOI: 10.1111/j.1749-4877.2010.00192.x.

Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C., Fortelius, M. (2010b). Precipitation and large herbivorous mammals, part II: Application to fossil data. *Evolutionary Ecology Research, 12*, 235-248.

Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C., Fortelius, M. (2010c). Precipitation and large herbivorous mammals I: estimates from present-day communities. *Evolutionary Ecology Research, 12*, 217–233.

Faith, J., Lyman, R. (2019). *Paleozoology and Paleoenvironments: Fundamentals, Assumptions, Techniques*. Cambridge, UK: Cambridge University Press.

Faith, J.T., Du, A., Rowan, J. (2019). Addressing the effects of sampling on ecometric-based paleoenvironmental reconstructions. *Palaeogeography, Palaeoclimatology, Palaeoecology, 528,* 175–185. DOI:10.1016/j.palaeo.2019.05.006.

Faith, J.T., Du, A., Behrensmeyer, A.K., Davies, B., Patterson, D.B., Rowan, J., Wood, B. (2021). Rethinking the ecological drivers of hominin evolution. *Trends in Ecology & Evolution, 36(9)*, 797–807. DOI: 10.1016/j.tree.2021.04.011.

Falchetti, E., Kingdon, J. (2013). *Kobus megaceros* Nile Lechwe (Mrs Gray's Lechwe). In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.455-460). London, United Kingdom: Bloomsbury Publishing.

Feakins, S.J., Levin, N.E., Liddy, H.M., Sieracki, A., Eglinton, T.I., Bonne!lle, R. (2013). Northeast African vegetation change over 12 m.y. *Geology, 41*, 295-298. DOI: 10.1130/G33845.1

Feibel, C.S. (1988). Paleoenvironments of the Koobi Fora Formation, Turkana Basin, Northern Kenya. PhD Thesis. The University of Utah.

Feibel, C.S. (1993). Freshwater stingrays from the Plio-Pleistocene of the Turkana Basin. *Lethaia, 26,* 359-366. DOI: 10.1111/j.1502-3931.1993.tb01542.x

Feibel, C.S. (2003). Stratigraphy and depositional setting of the Pliocene Kanapoi Formation, lower Kerio Valley, Kenya. In Harris, J.M., Leakey, M.G. (Eds*.). Geology and Vertebrate Paleontology of the Early Pliocene Site of Kanapoi, Northern Kenya* (pp. 9-20). Los Angeles, CA: Natural History Museum of Los Angeles County.

Feibel, C.S. (2011). A geological history of the Turkana Basin. *Evolutionary Anthropology, 20,* 206-216. DOI: 10.1002/evan.20331.

Feibel, C.S., Brown, F.H., McDougall, I. (1989). Stratigraphic context of fossil hominids from the Omo Group deposits: northern Turkana Basin, Kenya and Ethiopia. *American Journal of Physical Anthropology, 78,* 595-622. DOI: 10.1002/ajpa.1330780412

Feibel, C.S., Harris, J.M., Brown F.H. (1991). Palaeoenvironmental context for the late Neogene of the Turkana basin. In: Harris, J.H. (Ed.). *Koobi Fora Research Project, vol. 3. The fossil ungulates: geology, fossil artiodactyls and palaeoenvironments* (pp. 321-370). Clarendon Press, Oxford.

Feldesman, M.R. (2002). Classification trees as an alternative to linear discriminant analysis. *American Journal of Physical Anthropology, 119*, 257 – 275. DOI:10. 1002/ajpa.10102.

Fernández, M.H, Vrba, E.S. (2006). Plio-Pleistocene climatic change in the Turkana Basin (East Africa): Evidence from large mammal faunas. *Journal of Human Evolution, 50,* 595-626. DOI: 10.1016/j.jhevol.2005.11.004

Fick, S.E., Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37,* 4302–4315. DOI: 10.1002/joc.5086

Figueirido, B., Palmqvist, P., Pérez-Claros, J.A. (2009). Ecomorphological Correlates of Craniodental Variation in Bears and Paleobiological Implications for Extinct Taxa: An Approach Based on Geometric Morphometrics. *Journal of Zoology, 277*, 70–80. DOI: 10.1111/j.1469-7998.2008.00511.x

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Human Genetics, 7,* 179–188. DOI: 10.1111/j.1469-1809.1936.tb02137.x

Fleagle, J.G., Leakey, M. (2011). The Turkana basin. *Evolutionary Anthropology, 20*, 201. DOI: 10.1002/evan.20333

Forrest, F., Plummer, T., Raaum, R. (2018). Ecomorphological analysis of bovid mandibles from Laetoli Tanzania using 3D geometric morphometrics: Implications for hominin paleoenvironmental reconstruction. *Journal of Human Evolution, 114*, 20-34. DOI: 10.1016/j.jhevol.2017.09.010.

Fortelius, M., Eronen, J., Jernvall, J., Liu, L., Pushkina, D., Rinne, J., Tesakov, A., Vislobokova, I., Zhang, Z., Zhou, L. (2002). Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evolutionary Ecology Research, 4,* 1005–1016. DOI: 10.1016/j.palaeo.2006.03.042

Fortelius, M., Žliobaitė, I., Kaya, F., Bibi, F., Bobe, R., Leakey, L., Leakey, M., Patterson, D., Rannikko, J., Werdelin, L. (2016). An ecometric analysis of the fossil mammal record of the Turkana Basin. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1698), 20150232. DOI: 10.1098/rstb.2015.0232.

Fraser, D. Theodor, J. (2011). Comparing Ungulate Dietary Proxies Using Discriminant Function Analysis. *Journal of morphology, 272,* 1513-26. DOI: 10.1002/jmor.11001.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment, 83*, 287-302. DOI: 10.1016/S0034-4257(02)00078-0

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Wuang, *X.* (2010). MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*, 168–182. DOI: 10.1016/j.rse.2009.08.016

Frit, H., Bourgarel, M. (2013). *Aepyceros melampus* Impala. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.480-488). London, United Kingdom: Bloomsbury Publishing.

Garnier, S., Ross, N., Rudis, R., Camargo, A.P., Sciaini, M., Scherer, C. (2021). Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.1. DOI: https://sjmgarnier.github.io/viridis/

Geen, R., Bordoni, S., Battisti, D. S., Hui, K. (2020). Monsoons, ITCZs and the Concept of the Global Monsoon. *Reviews of Geophysics, 58*, e2020RG000700. DOI:10.1029/2020rg000700

Geraads, D., Bobe, R. (2020). Ruminants (Giraffidae and Bovidae) from Kanapoi. *Journal of Human Evolution, 140*, 102383. DOI: 10.1016/j.jhevol.2017.08.006

Geraads, D., Bobe, R., Manthi, F.K. (2013). New ruminants (Mammalia) from the Pliocene of Kanapoi, Kenya, and a revision of previous collections, with a note on the suidae. *Journal of African Earth Sciences, 85,* 53-61. DOI: 10.1016/j.jafrearsci.2013.04.006.

Giraud, T., Lambert, N. (2021). Cartography: Thematic Cartography. R package version 3.0.0. DOI: https://github.com/riatelab/cartography/

Good, S.P., Caylor, K.K. (2011). Climatological determinants of Woody cover in Africa. *PNAS, 108*, 4902–4907. DOI: 10.1073/pnas.1013100108.

Green, D.J., Gordon, A.D., Richmond, B.G. (2007). Limb-size proportions in *Australopithecus afarensis* and *Australopithecus africanus. Journal of Human Evolution, 52*, 187-200. DOI: 10.1016/j.jhevol.2006.09.001

Groves, C. (2013). FAMILY BOVIDAE Bovines, Antilopines. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.120-121). London, United Kingdom: Bloomsbury Publishing.

Guan, K., Wood, E.F., Medvigy, D., Kimball, J, Pan, M., Caylor, K.K., Sheffield, J., Xu, X., Jones, M.O. (2014). Terrestrial hydrological controls on land surface phenology of African savannas and woodlands. *Journal of Geophysical Research, 119,* 1652–1669. DOI: 10.1002/2013JG002572

Hagemann, S.I. (2010). Paleoecology and taphonomy of a hominid-bearing site: locality 261-1, Allia Bay, Kenya. M.Sc. Thesis. Rutgers University.

Haile-Selassie, Y., Woldegabriel, G., White, T.D., Bernor, R.L., Degusta, D., Renne, P.R., Hart, W.K., Vrba, E., Stanley, A., Howell, F.C. (2004). Mio-Pliocene mammals from the middle awash, Ethiopia. *Geobios, 37,* 536–552. DOI: 10.1016/j.geobios.2003.03.012.

Haile-Selassie, Y., Melillo, S. M., Vazzana, A., Benazzi, S., Ryan, T. M. (2019). A 3.8-million-year-old hominin cranium from Woranso-Mille, Ethiopia. *Nature*, *573*(7773), 214–219. DOI: 10.1038/s41586-019-1513-8.

Hammer, Ø., Harper, D.A.T. (2006). *Paleontological Data Analysis*. Oxford, UK: Blackwell Publishing Ltd.

Hammer, Ø., Harper, D.A.T., Ryan, P.D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica, 4(1)*, 9pp. DOI: palaeo-electronica.org/2001_1/past/issue1_01.htm.

Hanna, M. (2021). CAST: 'caret' Applications for Spatial-Temporal Models.R package version 0.5.1. DOI: https://CRAN.R-project.org/package=CAST

Hanot, P., Guintard, C., Lepetz, S., Cornette, R. (2017). Identifying domestic horses, donkeys and hybrids from archaeological deposits: a 3D morphological investigation on skeletons. *Journal of Archaeological Science, 78,* 88 – 98. DOI: 10.1016/j.jas. 2016.12.002.

Hardt, T., Hardt, B., Menke, P.R. (2007). Paleoecology: An Adequate Window on the Past? In Henke, W., Hardt, T., Tattersall, I. (Eds.). *Handbook of Paleoanthropology Volume I Methods and Approaches,* (pp. 503-554). Berlin, Germany: Springer-Verlag.

Hare, W. (2003). Assessment of Knowledge on Impacts of Climate Change, Contribution to the Specification of Art, 2 of the UNFCCC: Impacts on Ecosystems, Food Production, Water and Socio-economic Systems.

Harris, J.M. (1983). Background to the study of the Koobi Fora fossil fauna. In: Harris, J.M. (Ed.), *Koobi Fora Research Project: Volume II: The Fossil Ungulates, Proboscidea, Perissodactyla, and Suidae* (pp. 1-21.). Oxford University Press, Oxford.

Harris, J.M., Leakey, M.G. (2003). *Contributions in Science: Geology and Vertebrate Paleontology of the Early Pliocene Site of Kanapoi, Northern Kenya.* Los Angeles, CA: Natural History Museum of Los Angeles County.

Harris, J.M., Leakey, M.G., Cerling, T.E., Winkler, A.J. (2003). Early Pliocene tetrapod remains from Kanapoi, Lake Turkana Basin, Kenya. In Harris, J.M., Leakey, M.G. (Eds.). *Geology and Vertebrate Paleontology of the Early Pliocene Site of Kanapoi, Northern Kenya* (pp. 39-113). Los Angeles, CA: Natural History Museum of Los Angeles County.

Hastie, T., Tibshirani, R., Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd edition). New York, NY: Springer.

Hijmans, R. (2020). raster: Geographic Data Analysis and Modeling. R package version 3.4-5. DOI: https://CRAN.R-project.org/package=raster

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25,* 1965-1978. DOI: 10.1002/joc.1276

Hirota, M., Holmgren, M., Van Nes, E.H., Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science, 334,* 232–235. DOI: 10.1126/science.1210657

Ho, T.K. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 278–282. DOI: 10.1109/ICDAR.1995.598994

Hopley, P., Marshall, J.D., Weedon, G.P., Latham, A.G., Herries, A.I.R., Kuykendall, K.L. (2007). Orbital forcing and the spread of C4 grasses in the late Neogene: stable isotope evidence from South African speleothems. *Journal of Human Evolution, 53*, 620-634. DOI: 10.1016/j.jhevol.2007.03.007

Hosseini, M., Powell, M., Collins, J., Callahan-Flintoft, C., Jones, W., Bowman, H., Wyble, B. (2020). I tried a bunch of things: The dangers of unexpected overfitting in classification of Brain data. *Neuroscience & Biobehavioral Reviews, 119,* 456-467. DOI: 10.1016/j.neubiorev.2020.09.036

Hussain, S.T., Sondaar, P.Y., Shah, S.M.I., Thewissen, J.G.M., Cousin, E.F., Spoor, C.F. (1983). Fossil mammal bones of Pakistan, a field atlas. Part I: The artiodactyl astragalus. *Geological Survey of Pakistan, 14*, 1–15.

Iio, A., Hikosaka, K., Anten, N.P.R., Nakagawa, Y., Ito, A. (2013). Global dependence of field-observed leaf area index in woody species on climate: A systematic review. *Global Ecology and Biogeography.* DOI: 10.1111/geb.12133.

IUCN. (2015). The IUCN Red List of Threatened Species. Version 2015–4 [http:www.//iucnredlist.org/: downloaded 2 June 2015].

Janis, C.M. (1988). An estimation of tooth volume and hypsodonty indices in ungulate mammals, and the correlation of these factors with dietary preference. In: Russell, D.R., Santoro, J.P., Sigogneau-Russell, D. (Eds.), *Teeth Revisited, Proceedings of the VIIth International Symposium on Dental Morphology, Paris*, *1986. Mémoires du Museum national d'histoire naturelle, 53*, 367-387.

Janis, C.M., Fortelius, M. (1998). On the means whereby mammals achieve increased functional durabitily of their dentitions, with special reference to limiting factors. *Biological Reviews, 63,* 197-230. DOI: 10.1111/j.1469-185X.1988.tb00630.x.

Jarman, P.J. (1974). The social organisation of antelope in relation to their ecology. *Behaviour, 48*, 215–267. DOI:10.1163/156853974X00345

Jeffery, R., Nefdt, R. (2013). *Kobus leche* Southern Lechwe. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.449-454). London, United Kingdom: Bloomsbury Publishing.

Jenkins, R. (2013). *Kobus vardonii* Puku. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.445-449). London, United Kingdom: Bloomsbury Publishing.

Joordens, J.C.A., Feibel, C.S., Vonhof, H.B., Schulp, A.S., Kroon, D. (2019). Relevance of the eastern African coastal forest for early hominin biogeography. *Journal of Human Evolution, 131*, 176–202. DOI: 10.1016/j.jhevol.2019.03.0

Jordan, M.I., Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*, 255–260. DOI:10.1126/science.aaa8415.

Jungers W. (1991). Scaling of postcranial joint size in hominoid primates. *Human Evolution, 6*, 391–399. DOI:10.1007/BF02435532

Kappelman, J. (1988). Morphology and locomotor adaptations of the bovid femur in relation to habitat. *Journal of Morphology, 198*, 119–130. DOI: 10.1002/jmor.1051980111

Kappelman, J. (1991). The paleoenvironment of Kenyapithecus at Fort Ternan. *Journal of Human Evolution, 20*, 95–129. DOI: 10.1016/0047-2484(92)90072-H

Kappelman, J., Plummer, T.W., Bishop, L., Duncan, A., Appleton, S. (1997). Bovids as indicators of Plio-Pleistocene paleoenvironments in East Africa. *Journal of Human Evolution, 32,* 229–256. DOI: 10.1006/jhev.1996.0105

Kidwell, S.M., Behrensmeyer, A.K. (1988). Overview: Ecological and evolutionary implications of taphonomic processes. *Palaeogeography, Palaeoclimatology, Palaeoecology, 63(1-3)*, 1–13. DOI:10.1016/0031-0182(88)90087-9.

Kimura, T. (2019). How did humans acquire erect bipedal walking? *Anthropological Sciences, 127*, 1-12. DOI: 10.1537/ase.190219

Kingston, J.D. (2007). Shifting adaptive landscapes: Progress and challenges in reconstructing early hominid environments. *American Journal of Physical Anthropology, 134(S45)*, 20–58. DOI: 10.1002/ajpa.20733.

Kingston, J.D., Harrison, T. (2007). Isotopic dietary reconstructions of Pliocene herbivores at Laetoli: Implications for early hominin paleoecology. *Palaeogeography, Palaeoclimatology, Palaeoecology, 243,* 272-306. DOI: 10.1016/j.palaeo.2006.08.002.

Kingston, J.D., Marino, B.D., Hill, A. (1994). Isotopic evidence for Neogene hominid paleoenvironments in the Kenya Rift Valley. *Science, 264***,** 955–959. DOI: 10.1126/science.264.5161.955.

Kingston, J.D., Deino, A.L., Edgar, R.K., Hill, A. (2007). Astronomically forced climate change in the Kenyan Rift Valley 2.7e2.55 Ma: implications for the evolution of early hominin ecosystems. *Journal of Human Evolution, 53,* 487-503. DOI: 10.1016/j.jhevol.2006.12.007.

Kovarovic, K., Andrews, P. (2007). Bovid postcranial ecomorphological survey of the Laetoli paleoenvironment. *Journal of Human Evolution, 52,* 663–680. DOI: 10.1016/j.jhevol.2007.01.001

Kovarovic, K., Denise, S.F., Lintulaakso, K. (2018). Mammal Community Structure Analysis. In Croft, D., Simpson, S., Su, D. (Eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp. 351-372). Cham, Switzerland: Springer.

Kovarovic, K., Aiello, L.C., Cardini, A., Lockwood, C.A. (2011). Discriminant Function Analysis in archaeology: are classification rates too good to be true? *Journal of Archaeological Science, 38*, 3006-3018. DOI: 10.1016/j.jas.2011.06.028

Kovarovic, K., Faith, J.T., Jenkins, K.E., Tryon, C.A., Peppe, D.J. (2021). Ecomorphology and ecology of the grassland specialist, Rusingoryx atopocranion (Artiodactyla: Bovidae), from the late Pleistocene of western Kenya. *Quaternary Research, 101,* 187–204. DOI: 10.1017/qua.2020.102.

Kuhn, M. (2021). caret: Classification and Regression Training. R package version 6.0-88. DOI: https://CRAN.R-project.org/package=caret

Kuhn, M., Johnson, K. (2013). *Applied predictive modeling.* New York, NY: Springer.

Kullmer, O., Sandrock, O., Viola, T.B., Hujer, W., Said, H., Seidler, H. (2008). Suids, elephantoids, paleochronology, and paleoecology of the Pliocene hominid site Galili, Somali region, Ethiopia. *Palaios, 23*, 452-464. DOI: 10.2110/palo.2007.p07-028r

Laporte, L.F., Zihlman, A.L. (1983). Plates, Climates and Hominoid Evolution*. South African Journal of Science, 79*, 96–110.

Lawing, A.M., Polly, P.D. (2010). Geometric morphometrics: recent applications to the study of evolution and development*. Journal of Zoology, 280*, 1-7. DOI: 10.1111/j.1469-7998.2009.00620.x

Lawing, A.M., Head, J.J. and Polly, P.D. (2012). The ecology of morphology: the ecometrics of locomotion and macroenvironment in North American snakes. In Louys, J. (ed.). *Paleontology in Ecology and Conservation* (pp. 117–146). Berlin, Germany: Springer.

Leakey, M., Feibel, C., McDougall, I., Walker, A. (1995). New four-million-year-old hominid species from Kanapoi and Allia Bay, Kenya. *Nature, 376,* 565–571. DOI: 10.1038/376565a0.

Leakey, M.G., Feibel, C.S., McDougall, I., Ward, C.V., Walker, A. (1998). New specimens and confirmation of an early age for *Australopithecus anamensis*. *Nature, 393*, 62-66. DOI: 10.1038/29972.

Lee-Thorp, J., Sponheimer, M., Luyt, J. (2007). Tracking changing environ- ments using stable carbon isotopes in fossil tooth enamel: an example from the South African hominin sites. *Journal of Human Evolution, 53*, 595-601. DOI: 10.1016/j.jhevol.2006.11.020

Lepre, C.J., Quinn, R.L., Joordens, J.C.A., Swisher III, C.C., Feibel, C.S. (2007). Plio-Pleistocene facies environments from the KBS Member, Koobi Fora Formation: implications for climate controls on the development of lake-margin hominin habitats in the northeast Turkana Basin (northwest Kenya). *Journal of Human Evolution, 53*, 504-514. DOI: 10.1016/j.jhevol.2007.01.015

Levin, N.E. (2015). Environment and climate of early human evolution. *Annual Review of Earth and Planetary Science, 43*, 405-429. DOI: 10.1146/annurev-earth-060614-105310.

Li, S-F., Jacques, F.M.B., Spicer, R.A., Su, T., Spicer, T.E.V., Yang, J., Zhou, Z-K. (2016). Artificial neural networks reveal a high-resolution climatic signal in leaf physiognomy. *Palaeogeography, Palaeoclimatology, Palaeoecology, 442,* 1 – 11. DOI: 10.1016/j.palaeo.2015. 11.005.

Lin, Y., Jeon, Y. (2002). Random forests and adaptive nearest neighbors. *Journal of the American Statistical Association,* 101–474. DOI: 10.1198/016214505000001230

Liu, L., Puolamäki, K., Eronen, J.T., Ataabadi, M., Hernesniemi E., Fortelius M. (2012). Dental functional traits of mammals resolve productivity in terrestrial ecosystems past and present. *Proceedings of the Royal Society B, 279*, 2793–2799. DOI: 10.1098/rspb.2012.0211.

Loh W-Y. (2011). Classification and regression trees. *WIREs Data Mining and Knowledge Discovery, 1,* 14–23. DOI: 10.1002/widm.8.

Lovari, S. (1987). Evolutionary aspects of the biology of chamois "Rupicapra" spp (Bovidae; Caprinae). In Soma, H. (Ed.). *The Biology and Management of Capricornis and Related Mountain Antelopes* (pp.51-61). New York, NY: Croom Helm. Doi: 10.1007/978-94-011-8030-6.

Macho, Gabriele & Leakey, Meave & Williamson, D.K. & Jiang, Y. (2003). Palaeoenvironmental reconstruction: Evidence for seasonality at Allia Bay, Kenya, at 3.9 million years. Palaeogeography, *Palaeoclimatology, Palaeoecology, 199,* 17-30. DOI: 10.1016/S0031-0182(03)00483-8.

MacLatchy, L., DeSilva, J.M., Sanders, W.J., Wood, B. (2010). Hominini. In Werdelin, L., Sanders, W.S. (eds.). *Cenozoic mammals of Africa* (pp. 471–540). Berkeley, CA: University of California Press.

MacLeod, N. (2007). Automated taxon identification in systematics: theory, approaches and applications. Boca Raton, FL: CRC Press.

MacLeod, N. (2017). On the use of machine learning in morphometric analysis. In Lestrel, P.E. (Ed.). *Biological shape analysis: proceedings of the 4th international symposium* (pp. 134–171). Singapore: World Scientific.

Magill, C.R., Ashley, G.M., Freeman, K.H. (2013). Water, plants, and early human habitats in eastern Africa. *PNAS, 110,* 1175–80. DOI: 10.1073/pnas.1209405109

Manti, F.K., Cerling, T.E., Chits, K.L., Blumenthal, S.A. (2020a). Diets of mammalian fossil fauna from Kanapoi, northwestern Kenya. *Journal of Human Evolution, 140*, 102338. DOI: 10.1016/j.jhevol.2017.05.005.

Manti, F.K., Ward, C.V., Plavcan, J.M. (2020b). Introduction to Kanapoi: A Pliocene Hominin Site in West Turkana Kenya. *Journal of Human Evolution, 140,* 102718. DOI: 10.1016/j.jhevol.2019.102718.

Marcé-Nogué, J., Esteban-Trivigno, S.D., Püschel, T.A., Fortuny, J. (2017). The intervals method: a new approach to analyse finite element outputs using multivariate statistics. *PeerJ, 5*, e3793. DOI: 10.7717/peerj.3793.

Marean, C.W., Anderson, R.J., Bar-Matthews, M., Braun, K., Cawthra, H.C., Cowling, R.M., Engelbrecht, F., Esler, K.J., Fisher, E., Franklin, J., Hill, K., Janssen, M., Potts, A.J., Zahn, R. (2015). A New research strategy for integrating studies of paleoclimate, paleoenvironment, and paleoanthropology. *Evolutionary Anthropology, 24*, 62-72. DOI: 10.1002/evan.21443.

Mares, M.A., Willig, M.R. (1994). Inferring biome associations of Recent mammals from samples of temperate and tropical faunas: paleoecological considerations. *Historical Biology, 8,* 31–48. DOI: 10.1080/10292389409380470

Maslin, M.A., Christensen, B. (2007). Tectonics, orbital forcing, global climate change, and human evolution in Africa: introduction to the African paleoclimate special volume. *Journal of Human Evolution, 53(5),* 443-464. DOI: 10.1016/j.jhevol.2007.06.005.

Maslin, M.A., Trauth, M.H. (2009). Plio-Pleistocene East African pulsed climate variability and its influence on early human evolution. In Grine, F.E., Fleagle, J.G., Leakey, R.E. (Eds.). *The First Humans: Origin and Early Evolution of the Genus Homo* (pp. 151-158). Stony Brook, NY: Springer.

Maslin, M.A., Brierly, C.M., Milner, A.M., Shultz, S., Trauth, M.H., Wilson, K.E. (2014). East African climate pulses and early human evolution. *Quaternary Science Reviews, 101,* 1-17. DOI: 10.1016/j.quascirev.2014.06.012.

Mayaux, P., Bartholomé, E., Fritz, S., Belward, A. (2004). A new land-cover map of Africa for the year 2000. *Journal of Biogeography*, *31* (6), 861–877. DOI: 10.1111/j.1365-2699.2004.01073.x.

McDougall, I., Brown, F.H. (2008). Geochronology of the pre-KBS, Omo Group, Turkana Basin. *Journal of the Geological Society, 165,* 549-562. DOI: 10.1144/0016-76492006-170

McDougall, I., Brown, F.H., Fleagle, J.G. (2005). Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature, 433*, 733–36. DOI: 10.1038/nature03258.

Meloro, C., Kovarovic, K. (2013). Spatial and ecometric analyses of the plio-pleistocene large mammal communities of the Italian peninsula. *Journal of Biogeography, 40*, 1451–1462. DOI: 10.1111/jbi.12113.

Meloro, C., Elton, S., Louys, J., Bishop, L.C., Ditchfield, P. (2013). Cats in the forest: predicting habitat adaptations from humerus morphometry in extant and fossil Felidae (Carnivora). *Paleobiology, 39*, 323–344. DOI: 10.5061/dryad.s587t

Meloro, C., Caceres, N.C., Carotenuto, F., Sponchiado, J., Melo, G.L., Passaro, F., Raia, P. (2015). Chewing on the trees: constraints and adaptation in the evolution of the primate mandible. *Evolution, 69,* 1690-1700. DOI: 10.1111/evo.12694

Mendelssohn, H., Yom-Tov, Y., Groves, C.P. (1995). *Gazella dorcas*. *Mammalian Species, 491*, 1–6.

Miles, L., Newton, A.C., DeFries, R.S., Ravilious, C., May, I., Blyth, S., Kapos, V., Gordon, J.E. (2006). A global overview of the conservation status of tropical dry forests. *Journal of Biogeography, 33*, 491–505. DOI: 10.1111/j.1365-2699.2005.01424.x

Mitteroecker, P., Bookstein, F. (2011). Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. *Evolution Biology, 38*, 100 – 114. DOI: 10.1007/s11692-011-9109-8.

Moore, J. (1992). Savanna chimpanzee. In: Nishida, T., McGrew, W.C., Marler, P., Pickford, M., de Waal, F.B.M. (Eds.). *Topics in Primatology: Human Origins/Behavior, Ecology, and Conservation/Evolutionary Biology, Reproductive Endocrinology, and Virology* (pp. 99-119). Tokyo, Japan: University of Tokyo Press.

Muller, M. (1996). A novel classification of planar four-bar linkages and its application to the mechanical analysis of animal systems. *Philosophical Transactions of the Royal Society B, 351*, 689–720. DOI: 10.1098/rstb.1996.0065

Navega, D., Vicente, R., Vieira, D.N., Ross, A.H., Cunha, E. (2015). Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. *International Journal of Legal Medicine., 129*, 651 – 659. DOI: 10.1007/ s00414-014-1070-5

Negash, E.W., Alemseged, Z., Wynn, J.G., Bedaso, Z.K. (2015). Paleodietary reconstruction using stable isotopes and abundance analysis of bovids from the Shungura Formation of South Omo, Ethiopia. *Journal of Human Evolution, 88,* 127-136. DOI: 10.1016/j.jhevol.2015.07.009

Newby, J. (2013). *Addax nasomaculatus* Addax. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.566-570). London, United Kingdom: Bloomsbury Publishing.

Nicholson, S.E. (2000). The nature of rainfall variability over Africa on time scales of decades to millenia. *Global and Planetary Change, 26,* 137-158. DOI: 10.1016/S0921-8181(00)00040-0

Nowak, R.M. (1999). *Walker's Mammals of the World* (6th ed., pp. 1181–1183). Baltimore, USA: Johns Hopkins University Press.

Passey, B.H., Levin, N.H., Cerling, T.E., Brown, F.H., Eiler, J.M. (2010). High-temperature environments of human evolution in East Africa based on bond ordering in paleosol carbonates. *PNAS*, *107*, 11245–11249. DOI: 10.1073/pnas.1001824107

Patterson, D.B., Howells, WW. (1967). Hominid humeral fragment from early Pleistocene of northwestern Kenya. *Science, 156*, 64-66. DOI: 10.1126/science.156.3771.64.

Patterson, D.B., Braun, D.R., Behrensmeyer, A.K., Merritt, S., Zliobaite, I., Reeves, J.S., Wood, B.A., Fortelius, M., Bobe, R. (2017). Ecosystem evolution and hominin paleobiology at East Turkana, northern Kenya between 2.0 and 1.4Ma. *Palaeogeography, Palaeoclimatology, Palaeoecology, 481,* 1–13. DOI: 10.1016/j.palaeo.2017.05.001.

Peterson, L.E. (2009). K-nearest neighbor. *Scholarpedia, 4,* 1883. DOI: 10.4249/scholarpedia.1883.

Pham, H.N.A., Triantaphyllou, E. (2007). The impact of overfitting and overgeneralization on the classification accuracy in data mining. In Maimon, O., Rokach, L. (Eds.). *Soft computing for knowledge discovery and data mining, part 4* (pp. 391–431). New York, NY, USA: Springer.

Pickford, M., Senut. B. (2001). The geological and faunal context of Late Miocene hominid remains from Lukeino, Kenya. *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, 332*, 145–152. DOI: 10.1016/S1251-8050(01)01528

Plummer, T. (2004). Flaked stones and old bones: biological and cultural evolution at the dawn of technology. *American Journal of Physical Anthropology, 47,* 118–164. DOI: 10.1002/ajpa.20157.

Plummer, T.W., Bishop, L.C. (1994). Hominid paleoecology at Olduvai Gorge, Tanzania as indicated by antelope remains. *Journal of human evolution*, *27,* 47–75.

Plummer, T.W., Bishop, L.C. and Hertel, F. (2008). Habitat preference of extant African bovids based on astragalus morphology: operationalizing ecomorphology for palaeoenvironmental reconstruction. *Journal of Archaeological Science, 35*, 3016–3027. DOI: 10.1016/j.jas.2008.06.015

Plummer, T.W., Ditchfield, P.W., Bishop, L.C., Kingston, J.D., Ferraro, J.V., Braun, D.R., Hertel, F., Potts, R. (2009). Oldest evidence of toolmaking hominins in a grassland-dominated ecosystem. *PLoS one, 4(9)*, e7199. DOI: 10.1371/journal.pone.0007199.

Plummer, T., Ferraro, J., Louys, J., Hertel, F., Alemseged, Z., Bobe, R., Bishop, L. (2015). Bovid ecomorphology and hominin paleoenvironments of the Shungura Formation, lower Omo River Valley, Ethiopia. *Journal of human evolution, 88.* DOI: 10.1016/j.jhevol.2015.06.006.

Polly, P. (2010). Tiptoeing through the trophics: geographic variation in carnivoran locomotor ecomorphology in relation to environment. In Goswami, A., Friscia, A. (eds.). *Carnivoran Evolution: New Views on Phylogeny, Form, and Function* (pp. 347–410). Cambridge, UK: Cambridge University Press.

Polly, D., Head, J. (2015). Measuring earth-life transitions: ecometric analysis of functional traits from late cenozoic vertebrates. In Polly, P.D., Head, J.J., Fox, D.L. (eds.). *Earth-Life Transitions: Paleobiology in the Context of Earth System Evolution* (pp. 21-46). Baltimore, MD: The Paleontological Society.

Polly, P.D., Eronen, J.T., Fred, M., Dietl, G.P., Mosbrugger, V., Scheidegger, C., Frank, D.C., Damuth, J., Stenseth, N.C., Fortelius, M. (2011). History matters: ecometrics and integrative climate change biology. *Proceedings of the Royal Society B: Biological Sciences, 278*(1709), 1131–1140. DOI:10.1098/rspb.2010.2233.

Potts, R. (1996). Evolution and climate variability. *Science, 273,* 922-923. DOI: 10.1126/science.273.5277.922

Potts, R. (1998a). Environmental hypotheses of hominin evolution. *American Journal of Physical Anthropology*, *41(suppl. 27)*, 93–136. DOI: 10.1002/(sici)1096-8644(1998)107:27+<93::aid-ajpa5>3.0.co;2-x.

Potts, R. (1998b). Variability selection in hominid evolution. *Evolutionary Anthropology*, *7*, 81-96. DOI: 10.1002/(SICI)1520-6505(1998)7:3<81::AID-EVAN3>3.0.CO;2-A.

Potts, R. (2007). Environmental hypotheses of Pliocene human evolution. In Bobe, R., Alemseged, Z., Behrensmeyer, A.K. (Eds.). *Hominin Environments in the East African Pliocene: An Assessment of the Faunal Evidence* (pp. 25-49). Dordrecht, Netherlands: Springer.

Potts, R. (2013). Hominin evolution in settings of strong environmental variability. *Quaternary Science Reviews, 73*, 1-13. DOI: 10.1016/j.quascirev.2013.04.003.

Prell, W.L., Kutzbach, J.E. (1987). Monsoon variability over the past 150,000 years. *Journal of Geophysical Research, 92,* 8411-8425. DOI: 10.1029/JD092iD07p08411

Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A., Solomon, A.M. (1992). Special Paper: A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate. *Journal of Biogeography, 19(2),* 117. DOI:10.2307/2845499

Prins, H.H.T., Sinclair, A.R.E. (2013). *Syncerus caffer* African Buffalo. In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.125-136). London, United Kingdom: Bloomsbury Publishing.

Prost, J.H. (1980). Origin of bipedalism. *American Journal of Physical Anthropology, 52*, 175-189. DOI: 10.1002/ajpa.1330520204

Püschel, T.A., Marcé-Nogué, J., Gladman, J.T., Bobe, R., Sellers, W.I. (2018). Inferring locomotor behaviours in Miocene New World monkeys using finite element analysis, geometric morphometrics and machine-learning classification techniques applied to talar morphology. *Journal of The Royal Society Interface, 15*, 20180520. DOI: 10.1098/rsif.2018.0520

Püschel, T.A., Marcé-Nogué, J., Gladman, J., Patel, B.A., Almécija, S., Sellers, W.I. (2020). Getting Its Feet on the Ground: Elucidating Paralouatta's Semi-Terrestriality Using the Virtual Morpho-Functional Toolbox. *Frontiers in Earth Scienc,. 8,* 79. DOI: 10.3389/feart.2020.00079

Püschel, H.P., Bertrand, O.C., O'Reilly, J.E., Bobe, R., Püschel, T.A. (2021). Divergence-time estimates for hominins provide insight into encephalization and body mass trends in human evolution. *Nature Ecology & Evolution, 5,* 808–819. DOI: 10.1038/s41559-021-01431-1

Quinn, R.L., Lepre, C.J. (2020). Revisiting the pedogenic carbonate isotopes and paleoenvironmental interpretation of Kanapoi. *Journal of Human Evolution, 140,* 102549 DOI: 10.1016/j.jhevol.2018.11.005.

Ranjitsinh, M.K. (1989). *The Indian Blackbuck*. Dehradun, India: Natraj Publishers.

Reed, K.E. (1997). Early hominid evolution and ecological change through the African Plio-Pleistocene. *Journal of Human Evolution, 32*, 289-322. DOI: 10.1006/jhev.1996.0106

Reed, K.E., Fish, J.L. (2005). Tropical and temperate seasonal influences on human evolution. In: Brockman, D.K., van Schaik, C.P. (Eds.). *Seasonality in Primates: Studies of Living and Extinct Human and Non-Human Primates* (pp. 491-520). Cambridge, UK: Cambridge University Press.

Reynolds, S.C., Wilkinson, D.M., Marston, C.G., O'Regan, H.J. (2015). The 'mosaic habitat' concept in human evolution: past and present. *Transactions of the Royal Society of South Africa, 70,* 57–69. DOI: 10.1080/0035919X.2015.1007490.

Ricklefs, R.E., Miles, D.B. (1994). Ecological and evolutionary inferences form morphology: an ecological perspective. In: Wainwright, P.C., Reilly, S.M. (eds). *Ecological morphology: integrative organismal biology*. Chicago, IL: University of Chicago Press.

Roche, D., Ségalen, L., Senut, B., Pickford, M. (2013). Stable isotope analyses of tooth enamel carbonate of large herbivores from the Tugen Hills deposits: Palaeoenvironmental context of the earliest Kenyan hominids. *Earth and Planetary Science Letters, 381*, 39–51. DOI: 10.1016/j.epsl.2013.08.021

Roth, J. (2004). "Bubalus bubalis" (On-line), Animal Diversity Web. Accessed August 09, 2021 at https://animaldiversity.org/accounts/Bubalus_bubalis/

RStudio Team. *(2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.*

Sanders, W.J. (2020). Proboscidea from Kanapoi, Kenya. *Journal of Human Evolution, 140,* 102547. DOI: 10.1016/j.jhevol.2018.10.013

Saylor, B.Z., Luis, G., Deino, A., Alene, M., Levin, N.E., Melillo, S.M., Peaple, M.D., Feakins, S.J., Bourel, B., Barboni, D., Novello, A., Sylvestre, F., Mertzman, S.A., Haile-Selassie, Y. (2019). Age and context of mid-Pliocene hominin cranium from Woranso-Mille, Ethiopia. *Nature, 573,* 220-224. DOI: 10.1038/s41586-019-1514-7

Schaeffer, B. (1947). Notes on the origin and function of the artiodactyl tarsus. *Am Mus Novi., 1356*, 1–24.

Schneider, T., Bischoff, T., Haug, G.H. (2014). Migrations and dynamics of the intertropical convergence zone. *Nature, 513,* 45–53. DOI: 10.1038/nature13636

Schoeninger, M.J., Reeser, H. (1999). Paleoenvironment of Allia Bay, East Turkana, Kenya 3·9 Ma: the stable isotope data. *Journal of Human Evolution, 33,* A21–A22.

Schoeninger, M.J., Reeser, H., Hallin, K. (2003). Paleoenvironment of *Australopithecus anamensis* at Allia Bay, East Turkana, Kenya: Evidence from mammalian herbivore enamel stable isotopes. *Archaeological and Anthropological Sciences, 22*, 200-207.

Scholz, C.A., Moore, T.C., Hutchinson, D.R., Golmshtok, A.J., Klitgord, K.D., Kurotchkin, A.G. (1998). Comparative sequence stratigraphy of low-latitude versus high-latitude lacustrine rift basins: seismic data examples from the East African and Baikal rifts. *Palaeogeography, Palaeoclimatology, Palaeoecology, 140*, 401-420. DOI: 10.1016/S0031-0182(98)00022-4

Scholz, C.A., Johnson, T.C., Cohen, A.S., King, J.W., Peck, J.A., Overpeck, J.T., Talbot, M.R., Brown, E.T., Kalindekafe, L., Amoako, P.Y.O., Lyons, R.P., Shanahan, T.M., Castañeda, I.S., Heil, C.W., Forman, S.L., McHargue, L.R., Beuning, K., Gomez, J., Pierson, J. (2007). East African megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. *PNAS, 104*, 16416–16421. DOI: 10.1073/pnas.0703874104

Scoon, R. (2018). *Geology of National Parks of Central/Southern Kenya and Northern Tanzania*. Cham, Switzerland: Springer.

Scott, R.S., Kappelman, J. and Kelley, J. (1999). The paleoenvironment of *Sivapithecus parvada. Journal of Human Evolution*, *36*, 245–274. DOI: 10.1006/jhev.1998.0269

Senut, B. (2015). Morphology and environment in some fossil Hominoids and Pedetids (Mammalia). *Journal of Anatomy, 228(4),* 700–715. DOI:10.1111/joa.12427.

Senut, B., Pickford, M., Gommery, D., Segalen, L. (2018). Palaeoenvironments and the origin of hominid bipedalism. *Historical Biology, 30*, 284-296. DOI: 10.1080/08912963.2017.1286337

Serber, G., Lee, A. (2003). *Linear Regression Analysis* (2nd edition). Hoboken, NJ: John Wiley & Sons, Inc.

Short, R.A., Lawing, A.M. (2021). Geography of artiodactyl locomotor morphology as an environmental predictor. *Diversity and Distributions, 27,* 1818–1831. DOI:10.1111/ddi.13371.

Sillero-Zubiri, C. (2013). Tragelaphus buxtoni Mountain Nyala (Gedemsa). In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.159-162). London, United Kingdom: Bloomsbury Publishing.

Siyum, Z.G. (2020). Tropical dry forest dynamics in the context of climate change: syntheses of drivers, gaps, and management perspectives. *Ecological Processes, 9(1).* DOI:10.1186/s13717-020-00229-6

Skelly, A.C., Dettori, J.R., Brodt, E.D. (2012). Assessing bias: the importance of considering confounding. *Evidence-based spine-care journal*, *3*(1), 9–12. DOI: 10.1055/s-0031-1298595.

Skinner, J.D. (2013). Antidorcas marsupialis Springbok (Springbuck). In Kingdon, J., Hoffmann, M. (Eds.). *Mammals of Africa: Volume VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (pp.398-403). London, United Kingdom: Bloomsbury Publishing.

Sonnenschein, A., VanderZee, D., Pitchers, W.R., Chari, S., Dworkin, I. (2015). An image database of Drosophila melanogaster wings for phenomic and biometric analysis. *GigaScience, 4,* 25. DOI: 10.1186/s13742- 015-0065-6.

Sponheimer, M., Loudon, J.E., Codron, D., Howells, M.E., Pruetz, J.D., Cordon, J., de Ruiter, D.J., Lee-Thorpe, J.A. (2006). Do "savanna" chimpanzees consume C4 resources? *Journal of Human Evolution, 51* (2), 128-133. DOI: 10.1016/j.jhevol.2006.02.002

Sponheimer, M., Alemseged, Z., Cerling, T.E., Grine, F.E., Kimbel, W.H., Leakey, M.G., Lee-Thorp, J.A., Manthi, F.K., Reed, K.E., Wood, B.A., Wynn, J.G. (2013). Isotopic evidence of early hominin diets. *PNAS, 110,* 10513–10518. DOI: 10.1073/pnas.1222579110.

Staver, A.C., Archibald, S., Levin, S. (2011). Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states. *Ecology, 92*, 1063–1072. DOI: 10.1890/10-1684.1

Stewart, K.M., Rufolo, S.J. (2020). Kanapoi revisited: Paleoecological and biogeographical inferences from the fossil fish. *Journal of Human Evolution, 140*, 102452. DOI: 10.1016/j.jhevol.2018.01.008

Tarca, A.L., Carey, V.J., Chen, X., Romero, R., Dr˘aghici, S. (2007). Machine learning and its applications to biology. *PLoS Computational Biology, 3,* e116. DOI: 10.1371/ journal.pcbi.0030116.

Teaford, M.F., Ungar, P.S. (2000). Diet and the evolution of the earliest human ancestors. *PNAS, 97*, 13506-13511. DOI: 10.1073/pnas.260368897

Tiercelin, J.J. (1990). Rift-basin sedimentation: responses to climate, tectonism and volcanism. Examples of the East African Rift. *Journal of African Earth Sciences, 10*, 283-305. DOI: 10.1016/0899-5362(90)90061-I

Traiser, C., Klotz, S., Uhl, D., Mosbrugger, V. (2005). Environmental signals from leaves—a physiognomic analysis of European vegetation. *New Phytologist, 166*(2), 465–484. DOI: 10.1111/j.1469-8137.2005.01316.x.

Trauth, M.H., Maslin, M.A., Deino, A., Strecker, M.R. (2005). Late Cenozoic moisture history of East Africa. *Science, 309*, 2051–2053. DOI: 10.1126/science.1112964

Trauth, M.H., Maslin, M.A., Deino, A., Bergner, A.G.N., Duhnforth, M., Strecker, M.R., (2007). High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. *Journal of Human Evolution 53*, 475–486. DOI: 10.1016/j.jhevol.2006.12.009

Uno, K.T., Polissar, P.J., Jackson, K.E., deMenocal, P.B. (2016). Neogene biomarker record of vegetation change in eastern Africa. *PNAS*, *113*, 6355–6363. DOI: 10.1073/pnas.1521267113.

van Bocxlaer, B., Schultheiß, R. (2010). Comparison of morphometric techniques for shapes with few homologous landmarks based on machine-learning approaches to biological discrimination. *Paleobiology, 36*, 497 – 515. DOI: 10.1666/08068.1.

van den Brink, V., Bokma, F. (2011). Morphometric shape analysis using learning vector quantization neural networks—an example distinguishing two microtine vole species. *Annales Zoologici Fennici, 48,* 359 – 364. DOI: 10.2307/23737098.

van der Klaauw, Cornelius J. (1948). Ecological Studies and Reviews: IV. Ecological Morphology. *Bibliotheca Biotheoretica, 4,* 27–111.

van der Made, J. (2011). Biogeography and climatic change as a context to human dispersal out of Africa and within Eurasia. *Quaternary Science Reviews, 30*, 1353-1367. DOI: 10.1016/j.quascirev.2010.02.028.

Van Valkenburgh, B. (1987). Skeletal indicators of locomotor behavior in living and extant carnivores. *Journal of Vertebrate Paleontology, 7,* 162–182.

Van Valkenburgh, B. (1988). Trophic Diversity in Past and Present Guilds of Large Predatory Mammals. *Paleobiology, 14,* 155–173.

Venables, W.N., Ripley, B.D. (2013). *Modern applied statistics with S-PLUS*. New York, NY: Springer.

Vermillion, W.A., Polly, P.D., Head, J.J., Eronen, J.T., Lawing, A.M. (2018). Ecometrics: a trait-based approach to paleoclimate and paleoenvironmental reconstruction. In Croft, D.A., Su, D.F., Simpson, S.W. (eds.). *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities* (pp. 373–394). Cham, Switzerland: Springer.

Vignaud, P., Duringer, P., Mackaye, H.T., Likius, A., Blondel, C., Boisserie, J-R., de Bonis, L., Eisenmann, V., Geraads, D., Guy, F., Lehmann, T., Lihoreau, F., Lopez-Martinez, N., Mourer-Chauviré, C., Otero, O., Rage, J-C., Schuster, M., Viriot, L., Zazzo, A., Brunet, M. (2002). Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. *Nature, 418*, 152–155. DOI: 10.1038/nature00880.

Villaseñor, A., Bobe, R., Behrensmeyer, A.K. (2020). Middle Pliocene hominin distribution patterns in Eastern Africa. *Journal of Human Evolution, 147*, 102856. DOI: 10.1016/j.jhevol.2020.102856.

Vrba, E.S. (1985). Environment and evolution: alternative causes of the temporal distribution of evolutionary events. *South African Journal of Science, 81*, 229-236.

Vrba, E.S. (1995a). On the connections between paleoclimate and evolution. In Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (eds.). *Paleoclimate and evolution with emphasis on human origins* (pp 24–45). New Haven, CT: Yale University Press.

Vrba, E.S. (1995b). The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate. In Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (eds.). *Paleoclimate and evolution with emphasis on human origins* (pp. 385–424). New Haven, CT: Yale University Press.

Vrba, E.S. (2007). Role of environmental stimuli in hominid origins. In Henke, W., Hardt, T., Tattersall, I. (Eds.). *Handbook of Paleoanthropology Volume III, Phylogeny of Hominids* (pp. 1441-1483). Berlin, Germany: Springer-Verlag.

Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (Eds.). (1995). *Paleoclimate and Evolution, with Emphasis on Human Origins*. New Haven, CT: Yale University Press

Ward, C., Leakey, M., Walker, A. (1999). The new hominid species *Australopithecus anamensis*. *Evolution Anthropology, 7,* 197-205. DOI: 10.1002/(SICI)1520-6505(1999)7:6<197::AID-EVAN4>3.0.CO;2-T.

Ward, C., Leakey, M., Walker, A. (2001). Morphology of *Australopithecus anamensis* from Kanapoi and Allia Bay, Kenya. *Journal of human evolution, 41*, 255-368. DOI: 10.1006/jhev.2001.0507.

Ward, C., Manthi, F., Plavcan, J. (2013). New fossils of *Australopithecus anamensis* from Kanapoi, West Turkana, Kenya. *Journal of Human Evolution, 65*, 501-524. DOI: 10.1016/j.jhevol.2013.05.006

Ward, C.V., Plavcan, J.M., Manthi, F.K. (2020). New fossils of *Australopithecus anamensis* from Kanapoi, West Turkana, Kenya (2012-2015). *Journal of human evolution*, *140*, 102368. DOI: 10.1016/j.jhevol.2017.07.008.

Webb, G.I., Boughton, J.R., Wang, Z. (2005). Not so naive Bayes: Aggregating onedependence estimators. *Machine Learning, 58,* 5–24. DOI: 10.1007/s10994-005- 4258-6.

Welch, B.L. (1939). Note on discriminant functions. *Biometrika, 31*, 218–220.

Werdelin, L., Sanders, W.S. (2010). *Cenozoic mammals of Africa*. Berkeley, CA: University of California Press.

Wheeler, P.E. (1991). The thermoregulatory advantages of hominid bipedalism in open equatorial environments: the contribution of increased convective heat loss and cutaneous evaporative cooling. *Journal of Human Evolution, 21*, 107–15. DOI: 10.1016/0047-2484(91)90002-D.

White, T.D., WoldeGabriel, G., Asfaw, B., Ambrose, S., Beyene, Y., Bernor, R.L., Boisserie, J.-R., Currie, B., Gilbert, H., Haile-Selassie, Y. (2006). Asa Issie, Aramis and the origin of *Australopithecus*. *Nature, 440*, 883-889. DOI: 10.1038/nature04629

White, T.D., Asfaw, B., Beyene, Y., Hailie-Selassie, Y., Lovejoy, C.O., Suwa, G., Woldegabriel, G. (2009a). Ardipithecus ramidus and the paleobiology of early hominids. *Science, 326*, 75–86. DOI: 10.1126/science.1175802.

White, T.D., Ambrose, S.H., Suwa, G., Su, D.F., DeGusta, D., Bernor, R.L., Boisserie, J., Brunet, M., Delson, E., Frost, S., Garcia, N., Giaourtsakis, J.X., Haile-Selassie, Y., Howell, F.C., Lehmann, T., Likius, A., Pehlevan, C., Saegusa, H., Semprebon, G., Teaford, M., Vrba, E. (2009b). Macrovertebrate paleontology and the Pliocene habitat of Ardipithecus ramidus. *Science, 326*, 87–93. DOI: 10.1126/science.1175822.

Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software, 21(12),* 1-20. DOI: 10.18637/jss.v021.i12.

Wickham, H., Henry, L. (2020). tidyr: Tidy Messy Data. R package version 1.0.2. https://CRAN.R-project.org/package=tidyr

Wickham, H., François, R., Henry, L., Müller, K. (2021). dplyr: AGrammar of Data Manipulation. R package version 1.0.5. DOI: https://CRAN.R-project.org/package=dplyr

Will, M., Krapp, M., Stock, J.T., Manica, A. (2021). Different environmental variables predict body and brain size evolution in Homo. *Nature Communications, 12*. Doi:10.1038/s41467-021-24290-7

Willmott CJ, Matsuura K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*(1), 79–82. DOI: 10.3354/cr030079.

Wilson, K.E., Maslin, M.A., Leng, M.J., Kingston, J.D., Deino, A.L., Edgar, R.K., Mackay, A.W. (2014). East African lake evidence for Pliocene millennial-scale climate variability. *Geology, 42*,955–58. DOI: 10.1130/G35915.1

WoldeGabriel, G., White, T.D., Suwa, G., Renne, P., de Heinzelin, J., Hart, W.K., Heiken, G. (1994). Ecological and temporal placement of early Pliocene hominids at Aramis, Ethiopia. *Nature, 371,* 330–333. DOI: 10.1038/371330a0

WoldeGabriel, G., Olago, D., Dindi, E., Owor, M. (2016). Genesis of the East African Rift System. In Schagerl M. (ed.). *Soda Lakes of East Africa* (pp. 25-59). Cham, Switzerland: Springer. DOI: 10.1007/978-3-319-28622-8_2.

Wolfe, J.A. (1995). Paleoclimatic estimates from tertiary leaf assemblages. *Annual Review of Earth and Planetary Sciences, 23,* 119–142. DOI: 10.1146/annurev.ea.23.050195.001003.

Wood, B. (2002). Hominid revelations from Chad. *Nature, 418* (6894), 133–135. DOI: 10.1038/418133a.

Wood, B., Strait, D. (2004). Patterns of resource use in early Homo and Paranthropus. *Journal of Human Evolution, 46*, 119–162. DOI: 10.1016/j.jhevol.2003.11.004.

Wright S. (1921). Correlation and causation. *Journal of Agricultural Research, 20,* 557–585.

Wynn, J.G. (2000). Paleosols, stable carbon isotopes and paleoenvironmental interpretation of Kanapoi, Northern Kenya. *Journal of Human Evolution, 39,* 411-432. DOI: 10.1006/jhev.2000.0431.

Wynn, J.G. (2004). Influence of Plio-Pleistocene aridification on human evolution. Evidence from paleosols of the Turkana Basin, Kenya. *American Journal of Physical Anthropology 123(2)*, 106-118. DOI: 10.1002/ajpa.10317.

Wynn, J.G., Reed, K.E., Sponheimer, M., Kimbel, W.H., Alemseged, Z., Bedaso, Z.K., Campisano, C.J. (2016). Dietary flexibility of *Australopithecus afarensis* in the face of palaeoecological change during the middle Pliocene: faunal evidence from Hadar, Ethiopia. *Journal of Human Evolution, 99*, 93–106. DOI: 10.1016/j.jhevol.2016.08.002.

Yan, X., Su, X.G. (2009). *Linear Regression Analysis: Theory and Computing*. London: UK: World Scientific Publishing Co.

Zachos, J.C., Pagini, M., Sloan, L., Thomas, E., Billups, K. (2001). Trends, rhythms and aberrations in global climate 65 Ma to Present. *Science 292*, 686–693. DOI: 10.1126/science.1059412.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., Fink, W.L., (2004). *Geometric Morphometrics for Biologists*: *A Primer.* San Diego, CA: Elsevier Inc.

Zhang G.P. (2008) Neural Networks For Data Mining. In: Maimon O., Rokach L. (eds). *Soft Computing for Knowledge Discovery and Data Mining.* Boston, MA: Springer DOI: 10.1007/978-0-387-69935-6_2.

Zhang, S., Wang, X., Hammarlund, E. U., Wang, H., Costa, M. M., Bjerrum, C. J., Connellyc, A.N., Zhanga, B., Bian, L., Canfield, D. E. (2015). Orbital forcing of climate 1.4 billion years ago. *PNAS,* 201502239. DOI:10.1073/pnas.1502239112

Žliobaitė, I., Rinne, J., Tóth, A. B., Mechenich, M., Liu, L., Behrensmeyer, A. K., Fortelius, M. (2016). Herbivore teeth predict climatic limits in Kenyan ecosystems. *PNAS, 113*(45), 12751–12756. DOI: 10.1073/pnas.1609409113.

Žliobaitė, I., Tang, H., Saarinen, J., Fortelius, M., Rinne, J., Rannikko, J. (2018). Dental ecometrics of tropical Africa: linking vegetation types and communities of large plant-eating mammals. *Evolutionary Ecology Research, 19*, 127–147.

Appendix A – Measurements of the sample used in the study

Table A1- Sample used in the ecomorphological analysis. Raw measurements (cm) from the fossil sample (Kanapoi and Allia Bay) taken from Dumouchel et al. (2021). Raw measurements (cm) from the extant sample taken from Barr (2014). Iso and gm transformations were performed by me.

No	Taxon	Tribe	Site	MML	WAF	B	ProxRad	Dist Rad	GM	MML.gm	WAF.gm	B.gm	ProxRad.gm	Dist Rad.gm	MML.iso	WAF.iso	B.iso	ProxRad.iso	Dist Rad.iso
KNMER39276	?	?	Allia Bay	3.6	2.4	1.65	1.05	0.75	1.621995518	2.219488254	1.479658836	1.01726545	0.647350741	0.462393386	0.09186375	-0.117367	0.1733525	-0.09430808	-0.1866964
AB-203	?	?	Allia Bay	3.6	2.1	2.1	0.55	0.5	1.342810587	2.68094400920583	1.56388400537007	1.56388400537007	0.409588668073113	0.372353334611921	0.2286199	-0.3192335	0.680916	-0.5497	-0.4001816
KNMER39277	?	?	Allia Bay	4.2	2.5	1.95	1.2	0.9	1.85750304	2.261099933	1.345892818	1.049796398	0.646028552	0.484521414	0.2174707	-0.3577818	0.2736706	-0.09904889	-0.1633626
KNMER39278	?	?	Allia Bay	3.7	2.45	1.7	1.15	0.8	1.699498992	2.177112206	1.441601325	1.000294797	0.67667001	0.470726963	0.0833518	-0.1452329	0.1776776	-0.02970325	-0.1656699
KNMER39286	?	?	Allia Bay	4.1	2.35	2.1	0.9	0.8	1.708758556	2.399402763	1.375267438	1.228962391	0.526698168	0.468176149	0.2613672	-0.4045246	0.4842396	-0.3521117	-0.2249413
KNMER39396	?	?	Allia Bay	3.3	2.2	1.8	0.7	0.9	1.524438593	2.164731341	1.443154227	1.180762549	0.459185436	0.590381275	0.0349356	-0.1429436	0.4256665	-0.3650212	0.02820542
KNMER39430	?	?	Allia Bay	4	2.45	2.15	1	0.95	1.820864497	2.196758741	1.345514729	1.180757823	0.549189685	0.521730201	0.1263505	-0.3296519	0.5195004	-0.2635337	-0.08429098
KNMER43069	?	?	Allia Bay	3.9	2.2	2	0.85	0.8	1.634576456	2.385939174	1.345914406	1.223558551	0.520012384	0.48942342	0.2571321	-0.4140477	0.4666412	-0.3382558	-0.1726707
KNMER43144	?	?	Allia Bay	4.05	2.4	1.85	1.2	1.05	1.866558938	2.169768078	1.285788491	0.991128628	0.642894245	0.562532465	0.187163	-0.371893	0.2240516	-0.06000681	0.01859604
KNMER43153	?	?	Allia Bay	3.85	2.35	1.95	1.2	0.95	1.822609421	2.112356029	1.289360174	1.069894612	0.658396684	0.521230708	0.1050843	-0.337275	0.3736871	-0.02154242	-0.04991815
KNMER43162	?	?	Allia Bay	4.2	2.6	1.95	1.25	1.1	1.964771227	2.137653454	1.323309281	0.992481961	0.636206385	0.559861619	0.1418103	-0.3120741	0.2418236	-0.07372832	0.01643552
KNMER43213	?	?	Allia Bay	2.2	1.3	1.45	0.2	0.1	0.60778896	3.61967747	2.13890032	2.38569651	0.32906159	0.16453079	0.1488947	-0.1718313	0.5866471	-0.4690437	-0.4476592
KNMER43221	?	?	Allia Bay	4.15	2.55	2	1.2	0.95	1.89018636	2.195550708	1.349073326	1.058096727	0.634858036	0.502595945	0.156627	-0.315563	0.3191063	-0.102586	-0.116258
KNMER43661	?	?	Allia Bay	3.15	1.8	1.6	0.9	0.8	1.455482901	2.164230166	1.236702952	1.099291513	0.618351476	0.549645757	0.1430825	-0.3577026	0.3343259	-0.08081709	-0.002867582
KNMKP56869	?	?	Kanapoi	6.7	4.85	2.5	2.1	2.05	3.22660603	2.076485303	1.503127421	0.774807949	0.650838677	0.635342518	0.03350235	0.06625735	-0.3060675	-0.07452418	0.2699994

Table A1 - Continued.

No	Taxon	Tribe	Site	MML	WAF	B	ProxRad	Dist Rad	GM	MML.gm	WAF.gm	B.gm	ProxRad.gm	Dist Rad.gm	MML.iso	WAF.iso	B.iso	ProxRad.iso	Dist Rad.iso
KNMKP56871	?	?	Kanapoi	3.5	2.05	1.6	1	0.7	1.517078263	2.307066212	1.351281638	1.05465884	0.659161775	0.461413242	0.2067861	-0.3131431	0.2138178	-0.07420324	-0.1793107
KNMKP56872	?	?	Kanapoi	6.3	4.6	2.55	1.85	1.9	3.040274369	2.072181401	1.51302134	0.838740091	0.608497713	0.624943597	-0.002848093	0.07720485	-0.1029998	-0.2059065	0.2170964
KNMKP58618	?	?	Kanapoi	4.15	2.6	1.95	1.05	0.95	1.838204361	2.257637991	1.414423801	1.060817851	0.571209612	0.516808697	0.1748617	-0.2524781	0.2767817	-0.246638	-0.1113892
KNMKP58638	?	?	Kanapoi	1.7	0.9	0.95	0.3	0.25	0.641941215	2.648217563	1.401997533	1.479886285	0.467332511	0.389443759	0.1435089	-0.2169062	0.2948406	-0.207707	-0.1655938
KNMKP58715	?	?	Kanapoi	7.25	5	3.4	2.4	1.6	3.427864374	2.115019501	1.458634139	0.991871214	0.700144387	0.466762924	0.04110392	-0.1729567	0.3656258	0.04855255	-0.3248247
KNMKP58719	?	?	Kanapoi	6.5	4.3	3	1.8	1.6	2.99625734	2.169373075	1.435123727	1.001249112	0.600749467	0.533999526	0.1555808	-0.2526257	0.329502	-0.2694664	-0.09400342
KNMKP59694	?	?	Kanapoi	6.4	4.55	3.25	1.8	1.5	3.030314467	2.111992029	1.501494333	1.072495952	0.593997758	0.494998132	-0.02888047	-0.06323341	0.5439506	-0.2970166	-0.2165552
KNMKP59695	?	?	Kanapoi	3.95	2.4	1.85	1.2	0.85	1.780393694	2.218610419	1.348016457	1.039096019	0.674008228	0.477422495	0.1692052	-0.3130212	0.2585849	-0.03324571	-0.1594981
KNMKP59794	?	?	Kanapoi	3.15	1.9	1.45	0.8	0.7	1.371904993	2.296077364	1.384935553	1.056924501	0.583130759	0.510239414	0.1701977	-0.2382453	0.1957392	-0.1719725	-0.09562765
KNMKP114	?	?	Kanapoi	4	2.5	1.95	1.3	0.8	1.825633474	2.191020299	1.369387687	1.068122396	0.712081597	0.43820406	0.1212697	-0.2832977	0.3173618	0.03480901	-0.2356476
KNMKP115	?	?	Kanapoi	4.2	2.7	2.05	1.1	0.95	1.892764759	2.218976226	1.426484717	1.083071729	0.58116044	0.501911289	0.1285088	-0.221619	0.3362247	-0.2280671	-0.1371161
KNMKP125	?	?	Kanapoi	3.45	2.2	2.1	0.6	0.55	1.3937824832509	3.01338268379137	1.57843854865262	1.50669134189569	0.430483240541625	0.394609637163156	0.1046028	-0.2005888	0.6918528	-0.4912247	-0.343244
KNMKP127	?	?	Kanapoi	4.7	2.8	2.55	1.3	1.1	2.168834542	2.167062498	1.291015956	1.175746674	0.599400265	0.50718484	0.1538517	-0.4622232	0.6364317	-0.182894	-0.1138528
KNMKP130	?	?	Kanapoi	2.4	1.5	1	0.4	0.6	0.97118675	2.471203402	1.544502126	1.029668084	0.411867234	0.617800851	0.1601872	-0.1072439	0.05721624	-0.3305976	0.001954639
KNMKP132	?	?	Kanapoi	4.5	3	2.1	1.1	1	1.9897071	2.261639414	1.507759609	1.055431727	0.55284519	0.502586536	0.1345337	-0.1325695	0.2624845	-0.3239579	-0.1656094
KNMKP36861	?	?	Kanapoi	3	1.75	1.35	0.95	0.75	1.382469594	2.170029644	1.265850626	0.97651334	0.687176054	0.542507411	0.1471912	-0.2971174	0.1491935	0.01945115	-0.0117195
KNMKP37377	?	?	Kanapoi	2.6	1.65	1.3	0.5	0.5	1.06873032	2.432793335	1.543888078	1.216396668	0.467844872	0.467844872	0.1249564	-0.1260408	0.2582027	-0.3073268	-0.1608536
KNMKP459	?	?	Kanapoi	4.5	2.95	1.85	1.6	1.2	2.161232538	2.082145221	1.364961867	0.855993035	0.740318301	0.555238726	0.09421588	-0.2115008	-0.004486132	0.1628909	0.02362544
KNMKP467	?	?	Kanapoi	4	2.5	1.8	1.25	0.8	1.782602458	2.243910291	1.402443932	1.009759631	0.701221966	0.448782058	0.1637599	-0.2528076	0.1852468	-0.001331261	-0.2243024
KNMKP494	?	?	Kanapoi	3.15	2	1.4	0.9	0.7	1.409164932	2.235366442	1.419280281	0.993496196	0.638676126	0.496748098	0.1256307	-0.1702257	0.1269801	-0.08650963	-0.1075273

Table A1 - Continued.

No	Taxon	Tribe	Site	MML	WAF	B	ProxRad	Dist Rad	GM	MML.gm	WAF.gm	B.gm	ProxRad.gm	Dist Rad.gm	MML.iso	WAF.iso	B.iso	ProxRad.iso	Dist Rad.iso
KNMKP498	?	?	Kanapoi	3.8	2.3	1.8	1	0.75	1.638207992	2.319607778	1.403973129	1.098761579	0.6104231	0.457817325	0.1982322	-0.284555	0.2839411	-0.1748494	-0.2116967
KNMKP539	?	?	Kanapoi	4.05	2.5	1.55	1.15	0.85	1.726508414	2.345774841	1.448009161	0.89776568	0.666084214	0.492323115	0.2536032	-0.2242168	-0.04798232	-0.08833488	-0.163664
KNMKP56870	?	?	Kanapoi	6.85	4.3	3.2	2.25	1.55	3.186872175	2.149442972	1.349285369	1.004119345	0.706021414	0.486370308	0.2020069	-0.4704641	0.4017215	0.08151177	-0.2250597
KNMKP58637	?	?	Kanapoi	3.5	2.4	1.95	1	0.8	1.672941586	2.092123258	1.434598805	1.165611529	0.597749502	0.478199602	-0.02060679	-0.1263156	0.4681034	-0.1483758	-0.1400261
KNMKP58653	?	?	Kanapoi	3.7	2.35	1.7	1.25	0.9	1.754584167	2.108761762	1.339348687	0.968890539	0.712419514	0.51294205	0.09000421	-0.2404593	0.1804777	0.07246667	-0.06389362
KNMKP58686	?	?	Kanapoi	4.5	2.65	2.6	1.15	0.85	1.978379943	2.274588366	1.339479815	1.314206611	0.581283693	0.429644469	0.1689944	-0.4578412	0.7769897	-0.2627173	-0.3064082
KNMKP58687	?	?	Kanapoi	4.35	2.8	2.8	1.3	1.15	2.195276874	1.981526818	1.275465538	1.275465538	0.592180428	0.523851917	-0.07238035	-0.3734099	0.9385282	-0.1425225	-0.03080587
KNMKP59795	?	?	Kanapoi	3.9	2.4	1.8	0.95	0.95	1.723451685	2.262900686	1.392554268	1.044415701	0.551219398	0.551219398	0.1865349	-0.2647067	0.2369254	-0.2612836	-0.04152063
KNMKP59838	?	?	Kanapoi	4.75	2.7	2.3	1.2	0.95	2.019937874	2.351557472	1.336674774	1.138648881	0.594077677	0.470311494	0.3106329	-0.4855992	0.4313781	-0.2480634	-0.2353414
AMNH81690	*Aepyceros melampus*	Aepycerotini	Extant	3.439	2.139	1.516	1	0.922	1.593729518	2.157831652	1.342134895	0.951227911	0.627459044	0.578517239	0.1314621	-0.2344217	0.1237885	-0.07887553	0.0388647
AMNH82050	*Aepyceros melampus*	Aepycerotini	Extant	3.318	2.247	1.4	1.05	0.873	1.570949435	2.112098534	1.43034521	0.891180816	0.668385612	0.555714895	0.05811579	-0.09222637	0.02784696	-0.01333151	0.002588571
AMNH83534	*Aepyceros melampus*	Aepycerotini	Extant	3.694	2.207	1.722	1.065	0.919	1.688882743	2.187244801	1.306781071	1.019608974	0.6305944	0.544146717	0.1667009	-0.3241178	0.2372865	-0.08555875	-0.02281304
AMNH85150	*Aepyceros melampus*	Aepycerotini	Extant	3.409	2.255	1.527	1	0.885	1.597023001	2.134596683	1.412002206	0.956154044	0.626165058	0.554156076	0.07579598	-0.1368392	0.1239851	-0.08724748	-0.004988323
AMNH233038	*Alcelaphus buselaphus*	Alcelaphini	Extant	4.355	2.898	1.847	1.385	1.133	2.054218442	2.120027701	1.410755517	0.899125411	0.674222357	0.551547964	0.09447135	-0.1592684	0.05365491	-0.00472862	-0.004590355
AMNH34717	*Alcelaphus buselaphus*	Alcelaphini	Extant	3.595	2.404	1.547	1.091	0.979	1.701947749	2.112285763	1.412499298	0.908958575	0.641030255	0.5752233	0.07133592	-0.1245094	0.06381653	-0.05837305	0.03815754
AMNH34725	*Alcelaphus buselaphus*	Alcelaphini	Extant	4.771	3.137	1.848	1.614	1.353	2.27093888	2.100893178	1.381366988	0.813760342	0.71071926	0.595788822	0.1246727	-0.1971097	-0.1077358	0.0984289	0.1123987
AMNH82033	*Alcelaphus buselaphus*	Alcelaphini	Extant	4.23	2.87	1.597	1.426	1.212	2.018505165	2.095610193	1.421844268	0.791179546	0.706463389	0.600444339	0.08491521	-0.1044282	-0.1477524	0.07392765	0.1052339
AMNH82159	*Alcelaphus buselaphus*	Alcelaphini	Extant	4.606	3.141	1.813	1.5	1.313	2.201046065	2.092641346	1.427048734	0.823699253	0.681494142	0.596534539	0.07986158	-0.1068646	-0.09214576	0.02363295	0.10449
AMNH88406	*Beatragus hunteri*	Alcelaphini	Extant	3.969	2.792	1.507	1.332	1.15	1.912415989	2.075385283	1.459933412	0.788008471	0.696501184	0.601333604	0.03895062	-0.02812322	-0.1472395	0.05006939	0.1006499
AMNH88407	*Beatragus hunteri*	Alcelaphini	Extant	3.85	2.735	1.562	1.2	1.1	1.850693946	2.080300748	1.477824038	0.844007732	0.648405428	0.594371642	0.02325575	-0.01099357	-0.04875621	-0.04823383	0.07823305

Table A1 - Continued.

No	Taxon	Tribe	Site	MML	WAF	B	ProxRad	Dist Rad	GM	MML.gm	WAF.gm	B.gm	ProxRad.gm	Dist Rad.gm	MML.iso	WAF.iso	B.iso	ProxRad.iso	Dist Rad.iso
AMNH88408	*Beatragus hunteri*	Alcelaphini	Extant	3.883	2.777	1.576	1.215	1.11	1.870853359	2.07552344	1.484349367	0.842396328	0.649436255	0.593312135	0.01529163	0.001611416	-0.05199886	-0.04659579	0.07729535
AMNH81716	*Connochaetes gnou*	Alcelaphini	Extant	4.355	3.008	1.88	1.32	1.178	2.073139911	2.100678289	1.450939218	0.906837011	0.636715348	0.568220212	0.04974039	-0.08136643	0.06782672	-0.08431927	0.02846617
AMNH81720	*Connochaetes gnou*	Alcelaphini	Extant	4.48	2.973	1.874	1.388	1.25	2.124752739	2.108480633	1.399221634	0.881984979	0.653252482	0.588303748	0.1005263	-0.1696209	0.03058847	-0.04052693	0.0806505
AMNH81716	*Connochaetes gnou*	Alcelaphini	Extant	4.355	3.008	1.88	1.32	1.178	2.073139911	2.100678289	1.450939218	0.906837011	0.636715348	0.568220212	0.04974039	-0.08136643	0.06782672	-0.08431927	0.02846617
AMNH81720	*Connochaetes gnou*	Alcelaphini	Extant	4.48	2.973	1.874	1.388	1.25	2.124752739	2.108480633	1.399221634	0.881984979	0.653252482	0.588303748	0.1005263	-0.1696209	0.03058847	-0.04052693	0.0806505
AMNH81722	*Connochaetes gnou*	Alcelaphini	Extant	4.761	3.33	1.905	1.55	1.3	2.274377179	2.093320335	1.464137097	0.837591943	0.681505255	0.571585053	0.05185367	-0.04918729	-0.0771776	0.01393817	0.04262565
AMNH27824	*Connochaetes taurinus*	Alcelaphini	Extant	5.093	3.375	2.212	1.532	1.365	2.399301351	2.12270126	1.406659484	0.921935045	0.638519209	0.568915614	0.1137703	-0.1979936	0.1161387	-0.09215947	0.03551155
AMNH54133	*Connochaetes taurinus*	Alcelaphini	Extant	5.361	3.602	2.241	1.677	1.474	2.545964734	2.105685098	1.414787861	0.880216434	0.658689407	0.578955388	0.103932	-0.1703647	0.02819077	-0.03778671	0.0703268
AMNH83502	*Connochaetes taurinus*	Alcelaphini	Extant	5.158	3.414	2.055	1.734	1.4	2.447635604	2.107339831	1.394815468	0.839585761	0.708438788	0.571980567	0.1241459	-0.198191	-0.06385386	0.09202274	0.05592644
AMNH83503	*Connochaetes taurinus*	Alcelaphini	Extant	4.921	3.396	1.892	1.584	1.45	2.356204991	2.088527958	1.441300741	0.802986161	0.672267483	0.615396371	0.07652815	-0.08029412	-0.1471389	0.003796758	0.1564928
AMNH113781	*Damaliscus lunatus*	Alcelaphini	Extant	4.613	3.376	1.918	1.488	1.219	2.222123675	2.07594206	1.519267374	0.863138277	0.669629696	0.548574327	-0.01519024	0.05490514	-0.0301015	-0.02165502	-0.01675852
AMNH34729	*Damaliscus lunatus*	Alcelaphini	Extant	4.556	3.031	1.921	1.385	1.25	2.149865847	2.119201998	1.409855412	0.893544126	0.644226244	0.581431628	0.1035637	-0.1639774	0.04687701	-0.06732637	0.061169
AMNH34730	*Damaliscus lunatus*	Alcelaphini	Extant	4.872	3.241	2.027	1.579	1.278	2.301604026	2.116784618	1.408148388	0.880690152	0.686043291	0.555264931	0.1082629	-0.1773605	0.02184401	0.02513137	0.006049538
AMNH82035	*Damaliscus lunatus*	Alcelaphini	Extant	4.412	2.878	1.877	1.363	1.2	2.08052785	2.120615689	1.3833028	0.902174898	0.655122209	0.576776706	0.1161789	-0.2045936	0.06879957	-0.03824057	0.05298631
AMNH42953	*Damaliscus pygargus*	Alcelaphini	Extant	3.452	2.25	1.48	1.05	0.95	1.628866851	2.119264689	1.381328375	0.908607109	0.644619908	0.583227536	0.09029345	-0.162292	0.06498782	-0.04654463	0.05240131
AMNH81727	*Damaliscus pygargus*	Alcelaphini	Extant	3.655	2.364	1.506	1.124	1.05	1.726883688	2.116529345	1.36893991	0.872091161	0.650883443	0.608031686	0.1079318	-0.1813038	0.01296524	-0.03300718	0.1029085
AMNH81729	*Damaliscus pygargus*	Alcelaphini	Extant	3.551	2.279	1.481	1.107	0.96	1.663465438	2.134700198	1.37003147	0.890310051	0.665478209	0.577108474	0.1136198	-0.1875939	0.03413522	-0.01422838	0.04219593
AMNH81787	*Damaliscus pygargus*	Alcelaphini	Extant	3.544	2.277	1.449	1.1	1.014	1.671364425	2.120423259	1.362359977	0.866956349	0.658144917	0.60668995	0.1118448	-0.1858445	0.004334512	-0.01952406	0.09759103
AMNH17276	*Sigmoceros lichtensteinii*	Alcelaphini	Extant	4.763	3.394	1.844	1.609	1.331	2.296241106	2.074259531	1.478067783	0.803051559	0.70071039	0.579642964	0.02723501	-0.004288295	-0.1493819	0.06425551	0.06651828
AMNH216382	*Sigmoceros lichtensteinii*	Alcelaphini	Extant	4.863	3.127	2.031	1.482	1.35	2.281312653	2.131667482	1.370702081	0.890276919	0.649625994	0.591764569	0.1535665	-0.2523934	0.04870152	-0.05415551	0.09254896

Table A1 - Continued.

No	Taxon	Tribe	Site	MML	WAF	B	ProxRad	Dist Rad	GM	MML.gm	WAF.gm	B.gm	ProxRad.gm	Dist Rad.gm	MML.iso	WAF.iso	B.iso	ProxRad.iso	Dist Rad.iso
AMNH216383	*Sigmoceros lichtensteinii*	Alcelaphini	Extant	4.771	3.228	1.94	1.472	1.375	2.271491082	2.100382448	1.421092966	0.854064546	0.64803248	0.605329253	0.09066871	-0.1305102	-0.03004874	-0.05466274	0.1253195
AMNH81740	*Antidorcas marsupialis*	Antilopini	Extant	2.852	2.003	1.178	0.9	0.772	1.361341594	2.094992184	1.471342687	0.865322859	0.661112541	0.567087646	0.02555034	-0.02520258	-0.01171142	-0.02195084	0.01731857
AMNH81745	*Antidorcas marsupialis*	Antilopini	Extant	2.869	1.83	1.278	0.843	0.761	1.339013466	2.142622216	1.366677816	0.954434016	0.629567978	0.568328862	0.08885455	-0.1649756	0.107779	-0.06384702	0.01868211

Table A2 - Measurements of the extant sample used for the ecometric analysis taken from Barr (2017). Measurements are in millimeters.

id_no	binomial	B	DistRad	ProxRad	MML	WAF	astragGeomeans
10167	*Hippotragus equinus*	23.0717	15.9025	17.77	56.50475	36.32195	26.6254715700296
10170	*Hippotragus niger*	23.885225	14.44625	16.75	54.691875	37.20755	25.9472079356556
11033	*Kobus leche*	21.4340666666667	10.6057	12.065	43.96	26.5589	20.0026622490875
11034	*Kobus megaceros*	18.348375	10.1375	12.1725	40.621975	25.422875	18.7835982544904
11035	*Kobus ellipsiprymnus*	24.0728666666667	13.4174	15.92	54.1925	35.3253	25.0399153985792
11036	*Kobus kob*	18.94086	9.945	11.559	40.25814	25.10386	18.5568300806317
11037	*Kobus vardonii*	19.9568666666667	10.1266666666667	12.4583333333333	42.4948666666667	26.3434	19.498689469235
12142	*Litocranius walleri*	13.9578	8.1675	8.7025	30.80495	20.49605	14.4333839692812
12670	*Madoqua kirkii*	7.87556	3.997	4.393	16.30906	9.3169	7.31973624335545
14603	*Neotragus batesi*	6.82156666666667	3.05	3.56716666666667	13.4794666666667	8.19576666666667	6.06392655511721
14604	*Nesotragus moschatus*	7.28096666666667	3.45666666666667	4.08333333333333	14.7219	9.4347	6.77502133886385
15573	*Oryx gazella*	20.490075	13.22125	15.2025	49.0327	34.220375	23.3292259862071
15730	*Ourebia ourebi*	11.0299666666667	5.715	6.81136666666667	23.5032666666667	14.9466333333333	10.856735581973
1676	*Antidorcas marsupialis*	12.0885	7.5175	8.66125	28.219675	18.8644	13.3181617788108
19308	*Raphicerus campestris*	10.489575	4.95875	5.59375	20.8831	12.9612	9.53355042042113
19390	*Redunca arundinum*	15.9813333333333	9.26583333333333	10.1375	35.3698166666667	22.4892833333333	16.4212932674271
19391	*Redunca fulvorufula*	13.913975	7.31375	8.5325	29.677	18.80765	13.7114830628899
19392	*Redunca redunca*	16.34675	8.7525	9.69	34.60045	21.0142	15.8743658512696
21203	*Sylvicapra grimmia*	10.681975	5.455	6.57935	22.60175	14.7553	10.5037380529653
21251	*Syncerus caffer*	31.606225	18.54375	22.0375	71.827225	50.4255	34.1990753078805
22046	*Tragelaphus buxtoni*	25.88566	12.28	16.72	54.50714	34.8687	25.169616980881
22047	*Tragelaphus eurycerus*	27.1237666666667	13.8033333333333	17.9266666666667	58.5895333333333	37.3227333333333	27.1222303338096
22050	*Tragelaphus spekii*	19.637725	9.485575	12.155	41.101025	25.663225	18.8631725194113
22051	*Tragelaphus scriptus*	15.101	7.31166666666667	9.75666666666667	31.9878	20.0434666666667	14.7182492690363
22052	*Tragelaphus angasii*	18.0964	9.49	12.5266666666667	39.6865333333333	26.3809	18.6434094856785

Table A2 - Continued.

id_no	**binomial**	**B**	**DistRad**	**ProxRad**	**MML**	**WAF**	**astragGeomeans**
22053	*Tragelaphus imberbis*	19.6152	9.924	12.055	41.435175	26.3221	19.1261379830999
22054	*Tragelaphus strepsiceros*	23.6042	12.8275	16.45	52.3354	36.164625	24.8242047743509
22055	*Tragelaphus oryx*	29.9500666666667	19.5483333333333	21.7333333333333	71.137825	49.483125	33.90301442049
30208	*Damaliscus pygargus*	14.7904	9.935	10.95375	35.503725	22.924425	16.7284844733534
4139	*Cephalophus dorsalis*	11.431	5.84625	7.8025	24.93615	16.173	11.6028071700026
4141	*Cephalophus leucogaster*	11.8894	5.705	7.11166666666667	24.4682333333333	14.55	11.1421889061134
4143	*Philantomba monticola*	7.60795	3.482325	4.4	15.304075	9.077075	6.94813946628668
4144	*Cephalophus natalensis*	10.985675	5.59125	7.03875	23.4795	14.5155	10.8061354900229
4146	*Cephalophus nigrifrons*	11.4693	5.53	6.46	23.2956	13.820675	10.5696241276908
4150	Cephalophus silvicultor	17.1867666666667	8.87666666666667	11.8766666666667	36.4979	23.9607666666667	17.3772660142011
4152	Cephalophus weynsi	12.61565	5.787125	7.6775	25.84755	14.449475	11.5923838825832
44172	Tragelaphus derbianus	30.943025	18.50665	23.23125	72.034125	50.1729	34.3869873987596
5228	Connochaetes gnou	18.8596333333333	12.425	14.1916666666667	45.3214	31.0357	21.5777213937721
5229	Connochaetes taurinus	21.00005	14.22375	16.315	51.333325	34.46825	24.3852958770006
550	Aepyceros melampus	15.414775	8.995	10.2875	34.64995	22.123225	16.1346612178186
6234	Beatragus hunteri	15.4824666666667	11.2	12.4916666666667	39.0077666666667	27.6822	18.784777280023
6235	Damaliscus lunatus	19.357775	12.365475	14.53875	46.135	31.31435	21.8913955041274
63541	Nanger soemmerringii	13.27545	7.735	9.2	30.4647	19.23875	14.0816722116148
811	Alcelaphus buselaphus	17.3043	11.9796	14.033	43.11054	28.9011	20.5044886308852
8971	Nanger granti	14.212025	8.83045	9.89625	32.901575	21.82875	15.4906985397051
8982	Eudorcas thomsonii	10.33355	6.26125	7.2225	23.78395	15.4279	11.1387841396673

Appendix B - Categories assigned to extant data for ecomorphological analysis and their reference

Taxon	Tribe	Habitat	References	Dry or Wet	References
Aepyceros melampus	Aepycerotini	LC	Barr, 2014	D	Frit & Bourgarel, 2013
Alcelaphus buselaphus	Alcelaphini	O		D	Nowak, 1999
Beatragus hunteri	Alcelaphini	O		D	Groves, 2013
Connochaetes gnou	Alcelaphini	O		D	
Connochaetes taurinus	Alcelaphini	O		D	
Damaliscus lunatus	Alcelaphini	O		W	Duncan, 2013
Damaliscus pygargus	Alcelaphini	O		W	David & Lloyd, 2013
Sigmoceros lichtensteinii	Alcelaphini	O		D	Castelló, 2016 (p.535)
Antidorcas marsupialis	Antilopini	O		D	Skinner, 2013
Antilope cervicapra	Antilopini	O		D	Ranjitsinh, 1989
Eudorcas thomsonii	Antilopini	O		D	Groves, 2013
Gazella gazella	Antilopini	O		D	Mendelssohn et al. 1995
Litocranius walleri	Antilopini	LC		D	Groves, 2013
Madoqua kirkii	Madoquini	HC		D	
Nanger granti	Antilopini	O		D	
Nanger soemmerringii	Antilopini	O		D	
Neotragus batesi	Neotragini	F		W	
Neotragus moschatus	Neotragini	F		W	
Ourebia ourebi	Ourebiini	LC		D	
Raphicerus campestris	Raphicerini	LC		D	
Bubalus bubalis	Bovini	HC	Castelló, 2016 (p.601)	W	Roth, 2004
Syncerus caffer	Bovini	LC	Barr, 2014	W	Prins & Sinclair, 2013
Taurotragus derbianus	Tragelaphini	HC		W	Groves, 2013
Taurotragus oryx	Tragelaphini	LC		D	Castelló, 2016 (p.555)
Tragelaphus angasii	Tragelaphini	HC		D	Anderson, 2013
Tragelaphus buxtoni	Tragelaphini	F		W	Sillero-Zubiri, 2013
Tragelaphus eurycerus	Tragelaphini	F		W	Groves, 2013
Tragelaphus imberbis	Tragelaphini	HC		D	
Tragelaphus scriptus	Tragelaphini	F		W	
Tragelaphus spekii	Tragelaphini	HC		W	
Tragelaphus strepsiceros	Tragelaphini	HC		D	
Ammotragus lervia	Caprini	LC	Cassinello, 2013	D	Castelló, 2016 (p.303)
Rupicapra rupicapra	Caprini	LC	Lovari, 1987	W	Lovari, 1987
Cephalophus dorsalis	Cephalophini	F	Barr, 2014	W	Groves, 2013
Cephalophus dorsalis	Cephalophini	F		W	
Cephalophus leucogaster	Cephalophini	F		W	
Cephalophus natalensis	Cephalophini	F		W	
Cephalophus nigrifrons	Cephalophini	F		W	
Cephalophus silvicultor	Cephalophini	F		W	
Cephalophus weynsi	Cephalophini	F		W	
Philantomba monticola	Cephalophini	F		W	
Sylvicapra grimmia	Cephalophini	LC		W	
Addax nasomaculatus	Hippotragini	O		D	Newby, 2013
Hippotragus equinus	Hippotragini	O		D	Groves, 2013
Hippotragus niger	Hippotragini	LC		W	
Oryx gazella	Hippotragini	O		D	
Kobus ellipsiprymnus	Reduncini	HC		W	
Kobus kob	Reduncini	LC		D	
Kobus leche	Reduncini	HC		W	Jeffery & Nefdt, 2013
Kobus megaceros	Reduncini	HC		W	Falchetti & Kingdon, 2013
Kobus vardonii	Reduncini	HC		W	Jenkins, 2013
Redunca arundinum	Reduncini	LC		D	Groves, 2013
Redunca fulvorufula	Reduncini	LC		D	
Redunca redunca	Reduncini	LC		D	

Appendix C – Random Forest prediction classification of categories tested in ecomorphological analysis for each fossil individual

Table C1 - Random Forest prediction classification of habitat category for individual fossils from Allia Bay (KNMER and AB) and Kanapoi (KNMKP).

NO	F	HC	LC	O
KNMER39276	0.404	0.07	0.394	0.132
AB-203	0.694	0.036	0.26	0.01
KNMER39277	0.206	0.66	0.128	0.006
KNMER39278	0.078	0.806	0.11	0.006
KNMER39286	0.202	0.242	0.534	0.022
KNMER39396	0.076	0.014	0.892	0.018
KNMER39430	0.042	0.288	0.63	0.04
KNMER43069	0.254	0.12	0.608	0.018
KNMER43144	0.112	0.442	0.374	0.072
KNMER43153	0.236	0.496	0.26	0.008
KNMER43162	0.004	0.416	0.054	0.526
KNMER43213	0.8	0.068	0.13	0.002
KNMER43221	0.324	0.52	0.15	0.006
KNMER43661	0.32	0	0.556	0.124
KNMKP56869	0.008	0.184	0.712	0.096
KNMKP56871	0.774	0	0.21	0.016
KNMKP56872	0.064	0.156	0.622	0.158
KNMKP58618	0.08	0.17	0.686	0.064
KNMKP58638	0.608	0.326	0.066	0
KNMKP58715	0.012	0.73	0.254	0.004
KNMKP58719	0.168	0.05	0.742	0.04
KNMKP59694	0.128	0.332	0.518	0.022
KNMKP59695	0.138	0.694	0.166	0.002
KNMKP59794	0.348	0	0.598	0.054
KNMKP114	0.224	0.76	0.016	0
KNMKP115	0.076	0.202	0.676	0.046
KNMKP125	0.722	0.03	0.242	0.006
KNMKP127	0.24	0.75	0.006	0.004
KNMKP130	0.37	0.09	0.102	0.438
KNMKP132	0.006	0.326	0.606	0.062
KNMKP36861	0.036	0	0.264	0.7
KNMKP37377	0.712	0.048	0.238	0.002
KNMKP459	0	0.014	0	0.986
KNMKP467	0.18	0.804	0.014	0.002
KNMKP494	0.338	0	0.538	0.124
KNMKP498	0.328	0.094	0.558	0.02
KNMKP539	0.226	0.296	0.054	0.424
KNMKP56870	0.05	0.6	0.338	0.012
KNMKP58637	0.432	0.06	0.434	0.074
KNMKP58653	0.058	0.822	0.112	0.008
KNMKP58686	0.39	0.568	0.042	0
KNMKP58687	0.59	0.364	0.002	0.044
KNMKP59795	0.02	0.038	0.866	0.076
KNMKP59838	0.152	0.71	0.138	0

Table C2 - Random Forest prediction classification of 'wet and dry' category for individual fossils from Allia Bay (KNMER and AB) and Kanapoi (KNMKP).

NO	D	W
KNMER39276	0.552	0.448
AB-203	0.272	0.728
KNMER39277	0.318	0.682
KNMER39278	0.162	0.838
KNMER39286	0.578	0.422
KNMER39396	0.87	0.13
KNMER39430	0.63	0.37
KNMER43069	0.74	0.26
KNMER43144	0.364	0.636
KNMER43153	0.472	0.528
KNMER43162	0.672	0.328
KNMER43213	0.202	0.798
KNMER43221	0.57	0.43
KNMER43661	0.696	0.304
KNMKP56869	0.68	0.32
KNMKP56871	0.292	0.708
KNMKP56872	0.542	0.458
KNMKP58618	0.738	0.262
KNMKP58638	0.42	0.58
KNMKP58715	0.264	0.736
KNMKP58719	0.218	0.782
KNMKP59694	0.174	0.826
KNMKP59695	0.474	0.526
KNMKP59794	0.38	0.62
KNMKP114	0.28	0.72
KNMKP115	0.702	0.298
KNMKP125	0.234	0.766
KNMKP127	0.086	0.914
KNMKP130	0.67	0.33
KNMKP132	0.448	0.552
KNMKP36861	0.946	0.054
KNMKP37377	0.256	0.744
KNMKP459	0.762	0.238
KNMKP467	0.356	0.644
KNMKP494	0.532	0.468
KNMKP498	0.694	0.306
KNMKP539	0.414	0.586
KNMKP56870	0.194	0.806
KNMKP58637	0.752	0.248
KNMKP58653	0.114	0.886
KNMKP58686	0.072	0.928
KNMKP58687	0.154	0.846
KNMKP59795	0.95	0.05
KNMKP59838	0.12	0.88