Ca' Foscari
University
of Venice

Ph.D. in Environmental Sciences
35th Cycle

Ph.D. thesis

Reconstructing past anthropogenic impact on Italian Dark Earth by means of a multi-analytical approach

SSD: CHIM/01

Ph.D. Coordinator
ch. prof. Enrico Bertuzzo

Ph.D. Supervisor
prof. Dario Battistel

Co-supervisor
prof. Lavinia de Ferri

Ph.D. candidate
Mara Bortolini
865529

Declaration

These doctoral studies were conducted under the supervision of Prof. Dario Battistel. The work presented in this thesis is the result of original research carried out by myself, in collaboration with others, whilst enrolled in the Department of Environmental Sciences, Informatics and Satatistics as a candidate for the degree of Doctor of Philosophy in Environmental Sciences. This work has not been submitted for any other degree or award in any other university or educational establishment. Where any part of this thesis has previously been published in scientific journals this has been clearly stated. Where I have consulted the published work of others, this is always clearly attributed. Where I have quoted from the work of others, the source is always given. I have acknowledged all main sources of help.

This work is part of the INSIDE The Intimate Structure of the Italian Dark Earths project, which has received funding from Ca' Foscari University under the SPIN call for proposals 2018.

Mara Bortolini,

March 2023

Abstract

Dark Earths (DE) are soil deposits resulting from past settlements and human activities on natural soils. Those soils are referred to as cultural layers as they embed several clues of past domestic habits. The first DE were discovered in Amazonia, and by the 1970s, DE layers were excavated under European towns. Since the first studies, scholars focused on the remarkable abundance of black particles present in the soil matrix. Indeed, the origin of these charred particles and the properties derived from them (i.e., the strong fertility power) have been studied by chemists, soil scientists, and archeologists, highlighting the complexity of the DE formation. In this thesis, DE from Verona were characterized starting from the microscopic level of the layers up to the chemical and physical properties of the inclusions. The samples were radiocarbon dated and referred to the Medieval and Bronze Age. The soil morphology of the deposits was described, whereas PAHs and Raman parameters of the charred particles permitted us to classify the type of fire event and to identify the type of biomasses burned. The typology of vegetation and the level of microbic manipulation of the organic matter in soil were defined with the use of *n*-alkanes and miliacin. Fecal input in DE was assessed through classification models starting from a database of the fecal sterols in feces from humans and animals. The results permitted to outline that various domestic activities were involved in the formation of these layers, and being strictly anthropogenic, differences were evidenced even in a short-scale distance; indeed, the human impact was mainly related to soil exploitation for agricultural and livestock activities.

Contents

List of abbreviations

ACL: average chain length
AD: Anno Domini
ADE: Amazonian Dark Earth
ASE: accelerated solvent extractor
BC: black carbon
BP: before present
CDL: Castelar di Leppia
CRA: conventional radiocarbon age
D: diamond-like
DCM: dichloromethane
DE: Dark Earths
EF: enrichment factor
FWHM: full-width half maximum
G: graphitic-like
GALA: galactosan
GC-MS: gas chromatography – mass spectroscopy
HEX: *n*-Hexane
HW: high weight
L_a: crystallite lattice
LDA: linear discriminant analysis
LEVO: levoglucosan
LW: low weight
MANNO: mannosan
MeOH: methanol
MP-AES: microwave plasma – atomic emission spectroscopy
MW: medium weight
OEP: odd-even predominance
PAH: polycyclic aromatic hydrocarbons
PYR: pyrogenic
RF: random forest
ROC: recalcitrant organic carbon
SIM: selected ion monitoring
SOM: soil organic matter
SPM: vicolo San Pietro in Monastero
TIC: total inorganic carbon
TOC: total organic carbon
VP: via Pigna

Preface

This thesis is divided into five chapters. The first one (Introduction) contains a description of the Dark Earth soil layers and a state of the art of the studied conducted on the same kind of layers in Europe. This introductive chapter also contains the explanation of the proxies analyzed and their meaning. The second chapter (Materials and Methods) focuses on the description of the procedures and methods involved in the analyses. The chapters three, four, and five report the results and their discussion. These Results and Discussion chapters follow the structure: presentation of the data and type of approach used for the interpretation, discussion, and, lastly, a summary of the main conclusions pointed out.

The last section regards the conclusions of the thesis, divided into the main contents pointed out throughout the results and discussion chapters.

1 Introduction

This chapter will be focused on the description of the dark earth soil layers starting from the main features determined by scholars. The first part of this introductory chapter will also include a review of the studies on European dark earth layers conducted so far. This review will involve a description of the progress in the techniques employed by researchers from the first studies to recent times.
The second part of the chapter will be focused on the description of the proxies that were analyzed in this project, focusing on their possible interpretation from a chemical and physical point of view in an archaeological context. The latter will be provided through a discussion based on the published studies that applied the chemical analysis to samples of historical and archaeological interest.

1.1 Dark Earth layers

During the process of settlement, humans affect the environment in a deep and even irreversible way, through deforestation, tilling lands, and replacing the original vegetation with foreign introduced species. The most radical changes mainly occur within the borders of the human settlement itself. During the construction of villages and placement of household activities, specific deposits are developed, known as cultural layers (Alexandra Golyeva et al., 2016). This term is used in archaeology to define a different deposit in a profile, where a hiatus between natural soils can be highlighted (Peterson, 1974).
The term cultural layer is a general definition that is used mostly in Central and Eastern Europe to classify several soil layers in which the occurrence of anthropogenic materials are underlined (Krupski et al., 2017). In this context, the definition 'cultural layer' can be classified under the term 'anthrosol' as *"(...) soil that has been profoundly*

modified by human activities" (IUSS Working Group WRB., 2014). One of the most studied types of Anthrosols has been the Amazonian Dark Earth (ADE) or *Terra Petra de Indio*, for its hypothetical solid interest in sustainable agriculture in the tropics. Hence, despite being extremely important for global climate regulation and for global biodiversity, tropical rainforests are ecosystems characterized by nutrient-poor soils and by high turnover rates of organic matter. ADE is an ancient soil layer, dated up to 7,000 years ago. It was found as patches all around Amazon (Glaser, 2007) and could act as a model for sustainable agricultural practices and for long-term CO_2 sequestration into the terrestrial ecosystems. In fact, the better nutrient retention capacity of ADE than surrounding infertile soils, coupled with the soil organic matter (SOM) stocks, raised the interest of scholars and researchers (Glaser, 2007; Wiedner et al., 2015). The process pointed to be responsible for ADE formation is the intensive human occupation that enriched -voluntarily or involuntarily- the soil with massive amounts of charcoal, human excrement, and food waste. This amount of organic matter has been processed by fungi and bacteria, metabolizing and stabilizing SOM (Glaser, 2007; Glaser and Birk, 2012). However, recent studies overturned this hypothesis reporting the formation of this horizon to be natural: the results of trace elements, radiocarbon activity and microcharcoal, brought Silva et al., (2021) to suppose an exogenous source of material by fluvial transportation, rather than human influence on the Amazonian soils. Nevertheless, this hypothesis was strongly argued by the geoarchaeological community (Lombardo et al., 2022), highlighting some ambiguities in the hypothesis proposed by Silva and colleagues.

In Europe, soil horizons with features like ADE were found under both urban centers and rural areas in Europe. The term 'dark earth' was firstly coined by Norman and Reader in 1912 to define the archaeological layer found in London (Norman and Reader, 1912). This was a soil layer between the Roman and the Medieval levels, poorly stratified and black in colour. Since the last decades of 1900, the interest in this typology of soils has started to grow in Europe together with the development of urban archaeology with the pioneering studies of Macphail and colleagues in London, Belgium, France, and Italy (Brogiolo et al., 1988; Brogiolo, 1987; Cammas, 2004; David, 2004; Galinié, 2007; Macphail and Linderholm, 2004; Macphail, 1994, 1981). The common feature between all the layers investigated in the cited studies is the age of these soils. Indeed, they are related to the post-classical and early medieval period. The dating of these stratifications was historically made observing the layers

underneath and the inclusions found in the DE. These are typically rich in animal remains (such as bones and teeth) and frequently pottery fragments and metal slags (Brogiolo et al., 1988; Brogiolo, 1987; Cammas, 2004; David, 2004; Galinié, 2007; Macphail and Linderholm, 2004; Macphail, 1994, 1981).

In the field of urban archaeology the study of these deposits resulted in a *sensu stricto* definition of dark earth as a uniform dark colored layer that occur between stratified antique and modern age deposits in towns (Devos et al., 2020; Macphail, 1994; Macphail et al., 2003a). This traditional stance put the emphasis on the lack of knowledge on the development of urban centers during the passage between the Roman and Medieval period.

The progress of urban archaeology and the appearance of similar deposits related to cities outside the borders of the Roman empire (Heimdahl, 2005; Krupski et al., 2017; Prokof'eva et al., 2010; Wouters, 2016) and to different historical periods (Devos et al., 2009; Wouters et al., 2017) highlighted the need to find a more inclusive definition of DE. The numerous researches conducted since the end of the last century raised the knowledge of DE in different aspects: timing, location, and use of the soil. Indeed, it was demonstrated that DE can be related not only to the Medieval Age but even to prehistorical periods (Madgwick et al., 2012; Nicosia et al., 2011). Moreover, DE were excavated not only in correspondence with urban centers in post-Roman layers, but also in small *extra-murae* settlements (Macphail et al., 2003b). These aspects emphasize the necessity to include in the term 'dark earth' a wider class of soils with similar characteristics. As a result, scholars elaborated on a *sensu lato* definition (Devos et al., 2009; Yannick Devos et al., 2013a; Heimdahl, 2005; Macphail, 2003, 2010; Nicosia et al., 2017; Nicosia and Devos, 2014) used in European urban archaeology to designate thick, dark coloured, humic, and homogeneous units, often rich in anthropogenic remains (charcoal, ceramic, brick, bone, mortar, coprolites, slag, etc.), covering large surfaces regardless of their age or geographical location (Devos et al., 2020).

Moreover, the features that make these soils recognizable and unique with respect to underlaying layers lie in the radical change in the deposition style of DE. This aspect suggests that a drastic change in the use of soil would have occurred, and DE layers embed the information necessary to reconstruct the causes of this change. This succession of soil layers is related to the context of the continuity in the use of space in post-classical cities (Nicosia, 2018).

The term 'dark earth' is not present in any soil classification system such as World Reference Base FAO or Soil Taxonomy, but it is a strictly archaeological field term. This means that it is impossible to find guidelines pointing to chemical-physical characteristics to be used to unequivocally define a soil as ‘dark earth’. This aspect entails that a case-by-case approach should be pursued in the study of these layers. For this reason many authors tried to dissuade the scientific community from the use of that term in order to avoid any misunderstanding (Brogiolo et al., 1988). Moreover, some scholars suggested that the prolonged presence of human occupation in a site always gives rise to the development of dark colored soil stratification (Carver, 1987; Macphail et al., 2003a).

However, it has been possible to identify some characteristics that are typically found in DE, such as particular features observed in European DE resulted from the analysis of the thin sections that will be explored in further paragraphs. Carbonates, iron-manganese oxides and phosphates are the most typical parent minerals; calcite can derive from the diagenesis of the parent minerals or from the dissolution of ashes, plaster, and mortar used for buildings (Borderie et al., 2015; Nicosia, 2018; Nicosia et al., 2017). Secondary phosphates are well present and connected to the decomposition of vegetal material, feces, urine, bones, and ashes. Both carbonates and phosphates have also been interpreted as the outcome of reworking by bioturbation and digging or of processes such as dumping or nightsoiling (Devos et al., 2017; McGowan and Prangnell, 2006; Nicosia et al., 2017). In fact, the enrichment of nutrients in the soil could be the result of waste discard and/or manuring practices. One of the most distinctive elements of DE is the high degree of bioturbation from the activity of moles, rodents, earthworms, ants, and enchytraeids (Borderie et al., 2015; Meuser, 2010; Nicosia et al., 2017). The presence of stratification of alluvium material between DE strata is sometimes detected and can be indicative of periods of flooding and underwater conditions of the site (Nicosia, 2018; Nicosia et al., 2017). Another common feature of DE soils, as previously briefly reported, is the presence of rests from anthropic activities. These can be divided into three main groups: (i) remains from housekeeping activities and related to the domestic context (house sweeping, hearth functioning and maintenance, food preparation, construction or destruction debris, waste from manufacturing productions), (ii) evidence of urban cultivation (manuring, plant residues), and (iii) urban activities that resulted in a mechanical reworking and homogenization of the deposit (such as soil extraction, quarrying, pits, excavation of wall foundations, post holes) (Meuser, 2010; Nicosia et al., 2017).

The first studies on London's DE allowed to formulate two principal hypotheses on the formation of these layers. On one hand, some samples indicated a process of fast waste accumulation due to an increase of the inhabitants in the town with respect to the past. On the other hand, the evidence of soil backfilling was considered indicative of the presence of home gardens. The garden settling inside Roman towns has been interpreted as a sign of a urban and social crisis associated with population decline during the Middle Ages (Macphail, 1981). Further studies broadened the discussion supposing that DE could have been the result of the handling of organic material coming from the deterioration and collapse of wooden buildings. This hypothesis was firstly formulated by Yule (1990) during the researches conducted on the DE found in London. His idea was that right after the Roman occupation these layers were formed as the result of the deterioration of clay and timber buildings.

During urban excavation studies conducted in Italy, DE formation has been associated to the Medieval crisis of the urban tissue, with *spolios* of Roman buildings, rubble discharges and alluvial events. Various hypotheses have been proposed to explain the genesis of DE: (i) abandonment, (ii) presence of vegetable gardens, (iii) structures in perishable materials and (iv) accumulation of waste (Nicosia, 2018).

From the first studies on European DE, it is evident that their formation is the result of a complex relationship between human land use, settlement changes, and natural events. This implies that each cultural layer is the outcome of a unique sequence of events involving multiple drivers (Macphail, 1994; Nicosia, 2018; Nicosia et al., 2012a; Wiedner et al., 2015; Wouters et al., 2019). Besides their common features, the specific mechanism of DE formation is still debated and the need for a multidisciplinary approach, when studying cultural layers, has been remarked (Devos et al., 2017; Paetsch et al., 2017; Vrydaghs et al., 2019; Wiedner et al., 2015).

Most of the studies on DE deposits conducted so far are strictly connected to the development of urban archaeology and the scholars first analyzing them come from a geological-archaeological background. In this framework, the published researches are mainly based on archaeological and geomorphological methods, most of all micromorphology (Borderie et al., 2015; Macphail et al., 1990b). Micromorphology is the study of undisturbed soil blocks in the form of thin sections. From the observation of thin sections, it is possible to define the general petrographic and mineralogical textures, as well as to describe more local aspects from the study of the inclusions (natural or artificial). The inclusions in soil, such as the remains of manufacts, bones,

and food particles, as well as charred fragments, are important clues to describe the habits occurring in ancient times (Karkanas and Goldberg, 2007; Macphail et al., 1990a). Micromorphology can also inform about the processes occurred after the sedimentation, such as clay mobilization, redox of the constitutive elements, and bioturbation phenomena.

In Ferrara (Italy), Cremaschi and Nicosia (2010) reconstructed the major causes of DE formation dated back between the 8th and the 13th century AD. They assessed the anthropogenic accretion at the site as a result of the trampling of domestic deposits generated by the activity of fireplaces, reworking of deposits after the Po river flood with evidence of construction, destruction or refurbishing of buildings, and sediment accretion inside holes and cesspits. They also showed the sequence of natural alluvial and anthropogenic deposits. The stratigraphic information achieved was correlated with the observations made during the archaeological campaign, which indicated that buildings were mostly constituted by wooden houses that underwent several phases of rebuilding, with the settling of open ground floor (Cremaschi and Nicosia, 2010).

Another example of DE study in Italy was conducted by Nicosia and colleagues (Nicosia, 2018; Nicosia et al., 2012a) in Florence (by *Biblioteca Magliabechiana*), where they found deposits dated back between the 7th and the 11th century AD. As with Ferrara DE, this site lies close to the Arno river, meaning that various anthropogenic activities and natural processes were involved in its genesis. One of the most prominent human activities detected in the first phase of the DE formation was the dumping of waste material: the accumulation of household waste could have been a practice aimed at the production of manure for horticulture. The second phase is characterized by a faster sedimentation rate for the dumping of material to increase the level of the ground or to protect the area against Arno floods. The evidence of waste management is an indication of strong and active anthropogenic sedimentation and the adaptation of urban spaces. This archaeological implication, evinced by the results from micromorphological analysis, contrasts with the vision of DE as the result of the desertion of urban structures during the medieval period.

The Swiss site of Basel-Gasfabrik was interested by research conducted on the anthropogenic occupation layers found within the early Iron Age (around 2nd century BC). The aim of the study was the reconstruction of everyday life inside an Iron Age settlement and the related waste and resources management. The research focused on different structures identified within the site: occupation layers, pits, and ditches. From the micromorphological analysis of the thin sections, Brönnimann and collegues (2020)

were able to identify a constant accumulation of material that permitted the installation of a DE layer. The occupation layers are formed by the continuous accumulation of remains from domestic and economic activities. In parallel, the authors evidenced the presence of garden cultivation, determined by the presence of microcharcoal, ash, and coprolite pieces together with residues of herbivore dung and bird guano. The thin section observation associated the structures of the ditches to different uses, but most of all drainage of surface water within the village. The settlement pits were initially used as cellar or storage spots, and secondly filled with waste material from artisanal craft and/or domestic waste such as ceramic, bone fragments, ash, charcoal, and glass droplets (Brönnimann et al., 2020).

Since the early 2000s, the study of the micromorphology of soil has been coupled with other approaches: the introduction of the analysis of pollen and phytoliths allowed improving the reconstruction of the environmental outline. Additionally, the analyses of soil chemical characteristics (pH, carbon and nitrogen content, major element analysis, cation exchange capacity) have become a common approach to reconstruct the geochemical background, and an important support for the micromorphological results. In fact, cultural layers are generally enriched in organic matter, resulting in a high C content and C/N ratio. The pH is also affected by the presence of construction materials containing oxides and carbonates (mainly CaO and $CaCO_3$). These materials increase the pH of cultural soils compared to the natural horizon soil (Alexandrovskaya and Alexandrovskiy, 2005; A Golyeva et al., 2016; Mazurek et al., 2016). Although $CaCO_3$ can act as a pH buffer in soil, it can undergo leaching reactions under humid conditions (e.g., high rainfall events). The gradual loss of carbonates (and also ashes) leads to a depletion of nutrients in the soil, thus bringing to a relative enrichment of resistant biomass burning particles, phytoliths, and archaeological fragments (Nicosia, 2018). In addition, phosphorous has been widely considered one of the most important elements connected to the anthropogenic origin of urban soil and cultural layers. Thus, the presence of substantial concentrations of P has been associated with the addition of excrement and manure to the soil as a fertilizer, and/or to the dumping of ashes and waste, especially from food (bones, fish, meat, plants) (Devos, 2018; Holliday and Gartner, 2007; Mazurek et al., 2016; Migliavacca et al., 2013; Nicosia, 2018; Vittori Antisari et al., 2013; Zhang et al., 2001).

Dark buried soil horizons were identified in Warsaw in correspondence of the Odra River fluvial terrace and associated to the early medieval period (13^{th}-14^{th} century AD).

The analysis conducted by Krupski and colleagues (2017) were focused on the description and origin of this layers via micromorphological and soil analysis (pH, cation exchange capacity, organic carbon content). The layer was thick (50-60 cm) and without traces of human and animal excrements. This aspect allowed to exclude the hypothesis of farm-yard manure process. Nevertheless, the authors suggested the possibility of a periodical addition of household waste to the soil due to the presence of isolated phytoliths, bone fragments and charred material. Therefore, the origin of these layers was attributed to a particular cultivation technology called 'double- or triple-depth hand digging' which consists in a periodical digging followed by a logical and systematic remixing of the soil. This procedure included the addition of organic rich waste material to the soil (Krupski et al., 2017).

In the context of environmental reconstruction and palaeobotanical research, one of the most studied remains are phytoliths. Phytoliths can provide information about the plant species typical of a specific geographical area and a time period, as well as about their original use (Vrydaghs, Ball, & Devos, 2016 and references therein). Phytoliths are the result of metabolic mechanisms occurring inside the plant body and involving the polymerization of silicic acid ($Si(OH)_4$) in the form of opal ($SiO_2\ n(H_2O)$). This process can occur in different parts of the plant such as leaves, inflorescence bracts, fruits, seeds, woody tissues, and roots. From the studies of the different phytoliths morphotypes produced by plants it is possible to describe the taxa that were present in a site (Piperno, 2006).

Devos et al. (2009) integrated micromorphology with the analysis of phytoliths and soil chemical composition for DE deposits in Belgium, and particularly in Brussels. These DE layers were dated from the 13th century AD. In this study, Devos and colleagues identified different changes in land use: several plow horizons, enriched with manure and lime, and the hypothesis of a change in the use of the soil was supported by the evidence of a hearth structure carved out. The discovery of a successive addition of soil sod followed by pasture activities demonstrated the complex phenomena at the base of DE formation. Additionally, a growing anthropogenic input in the surface layers was suggested by the decreasing presence of inclusions with depth.

In the area of Brussels other three sites were investigated to draw the formative processes of DE structure. Devos et al. (2017) defined different sequences of events and anthropogenic activities together with syn- and post- depositional processes. The most conspicuous ones are related to agricultural activities and waste management. The

integrated study of DE from different periods and locations permitted to reconstruct the urban development of the area. Indeed, during the pre-urban period (10^{th}-11^{th} century AD) the inhabitants of the neighborhood were mainly located in small villages sustained by agriculture, the surroundings were likely used as grasslands or for pasture. Then, between the 12^{th} and the 15^{th} century, the pre-urban nuclei changed into more urbanized communities, mostly dedicated to artisanal activities. Additionally, the researchers stated the presence of a densely occupied area in which they associated the DE to a phase of accumulation of excremental waste.

In Lier (Belgium), three different DE layers spanning between the 11^{th} and 15^{th} centuries AD were studied by Wouters et al., (2017). They were able to define the human activities that occurred in succession at the site. (i) During the first period, intense agricultural activity and manure addiction to the soil appeared to have taken place. Despite the limited content of pollen and phytoliths, the textural features of the matrix corroborate the interpretation. (ii) During the second phase, DE appeared to be the result of the preparation of the surface for a transformation of the site into an urban place, evidenced by a clear interruption of the deposition of agricultural microremains. (iii) The third DE layer unit was associated with market activity inferred from the presence of increasing anthropogenic fragments and evidence of intense trampling. Scholars described the latter sequence as the result of the sum of chaotic events involving manufacturing, artisanal activities including installation of hearths, trade, traffic, and waste accumulation.

In Antwerp (Belgium), the study of Crabtree et al., (2017), focused on the dark earth of the historical burg of the city between 8^{th}-10^{th} century AD. The results demonstrated that the area was interested by artisanal and market activity during the early medieval period. Antwerp was indeed an important trading and manufacturing settlement (Wouters, 2020) and the remnants of bones, antler and metallic materials indicated the presence of craft-working area. In parallel, remains of wooden trackways and floors covered with ashes suggested the presence of houses. The study of Crabtree and colleagues is supported by the identification of species of insects indicating a strong relationship between the city and the hinterland areas. The cumulative presence of species usually found in stored grains and legumes, together with the ones typical of wild and cultivated Brassica plants, suggested the close interconnection between the city and the surrounding countryside. This hypothesis was further supported by the presence of beetles' species typical of wetland and grassland that likely covered the fields around the city.

DE soils from the same sites in Antwerp with the same age of the previous ones were studied by Devos et al., (2013) through micromorphological analysis. The results indicated that the main forming phenomena were the accumulation of sediment, homogenization and, lastly, humification. The sedimentation of local material together with waste from anthropogenic activities (i.e., charcoal, bones, pottery fragments, ashes), was followed by a process of transformation into humic-like materials like sod. The action of mesofauna, plant roots, and humans brought to a homogenization of the soil. The main human-related activities were identified in crop growing and pasture. In particular, digging and ploughing were associated to the high degree of homogenization of the layers.

The elemental concentration detected in DE from a hillfort in the Czech Republic was the key to understanding the past settlement activity in the site between the 8th and the 13th centuries AD (Asare et al., 2021). Asare et al. (2021) focused on the detection of elements in soil such as N, P, Ca, Mn, Fe, Al, Cu, Zn, Sr and Rb, in addition to measuring its pH. The pH, in fact, was considered as an indicator for the occurrence of intensive and long-term past settlement activities in the hillfort. The accumulation of these elements in soils would be the result of the deposition of biomass ash, domestic waste, and excrements, and was positively affected by the high proportion of fine sand/loam, silt, and clay particles in the soil, as well as the inclusion of charcoal.

Wiedner et al. (2015) conducted a deep research focused on the chemical and physical description of DE from northern Germany dated 10th-11th century AD. The research also laid the foundations for the determination of fecal waste input, with the determination of fecal sterols and stanols in soil samples. The authors compared the European and the Amazonian DEs, suggesting an anthropogenic formation of the soil supported by the input of organic waste, fecal material, and charred organic matter. From the relative abundances of fecal sterols and stanols, they were able to determine the fecal input derived from pigs, cows, and sheep and the relatively less abundant presence of human feces in the samples. Another important aspect inferred from the analysis of the elements, was the higher fertility and nutrient retention of DE with respect to a less anthropologically influenced soil from an adjacent site. The studied site was unequivocally described as a settlement soil with a possible influence of horticultural activity in the surrounding village.

The multidisciplinary approach for a deeper characterization of European DE was followed by Nicosia et al. (2019) in the site of Padua, dated back between the 7th and the 10th century AD. From the chemical point of view, the focus was on the study of

humic substances and therefore on the humification processes, as well as on the analysis of fecal sterols and stanols to distinguish between two types of anthropogenic deposits: trampled, domestic occupation deposits and units deriving from dumping of domestic waste. The first is described as household deposits and is polluted by fecal material of mixed contribution, probably derived from the coexistence of humans and animals in the same living space. The second unit is characterized by many events of dumping of domestic waste with variable content of fecal sterols and stanols. The authors rejected the hypothesis of the collapse of the Roman sewer during the post-Roman period, as evinced by the mere accumulation of solid rubbish without any *extra muros* forming DE soils. In this framework, they explained the waste accretion as the result of a change in the life habits after the collapse of a Roman regulatory system. However, the fast dumping of heterogeneous material could be also due to the accumulation of backfilling to raise the ground level, or to fill negative structures (i.e., holes, ditches).

What is evident from the studies conducted so far is the synergy between disciplines and analytical approaches. The multidisciplinary methods including archaeological and pedological knowledge, and physical and chemical characterization of the samples are at the base of the description of DE phenomena. From literature, it is possible to note the importance of the organic fraction included in the soil and its key role in understanding the successions of events and soil uses made in the past. Conversely, the understanding of the possible sources of the charcoal fragments widely found in DE is still missing.

The integration of the determination of new biomarkers to the research approaches conducted so far and the application of new analytical protocols in the study of cultural layers is one of the aims of this project. The next paragraphs are focused on the description of the meanings and importance of new proxies in the context of archaeological research.

1.2 Proxies

Proxy means literally ‘substitute’ and it indicates that the information is not directly given by a measurable signal. This entails that the dynamics that lead to the accumulation and preservation of the proxy must be known (Ruddiman, 2014).

Molecules identical or similar to the precursor which they derive from are called ‘molecular proxies’. Their stability, specificity and concentration in the samples are factors that make these molecules highly characteristic indicators of the source they come from (Mazurek et al., 2016; Oros and Simoneit, 2001).

For example, charcoal particles, as well as polycyclic aromatic hydrocarbons and levoglucosan, can be considered a proxy of biomass burning events, because of the exclusive or particular origin process (Bucheli et al., 2004; Kirchgeorg et al., 2014; Oros et al., 2006; Simoneit, 2002; Simoneit and Elias, 2000).

The following paragraphs will focus on the description of the proxies that were employed in this study.

1.2.1 Charcoal particles

The progress of human societies and technology is strictly connected to the familiarity, handling, and domestication of fire. This strong relation has been used by researchers to obtain information on the development of human settlements in relation to fire proxies (Cockerton et al., 2015; D’Anjou et al., 2012; Schüpbach et al., 2015; Zou et al., 2010, among many others). The base of this approach is the quantification of fire markers in the organic fraction of soils ad sediments.

Charcoal is the product of the incomplete combustion of organic matter (Whitlock and Larsen, 2002) and is composed of ordered and amorphous phases: the first one is responsible of the structural and chemical stability and is constituted by polyaromatic graphitic domains; the disordered structures are described by carboxylic and aliphatic groups (Cohen-Ofri et al., 2006; De Sousa et al., 2020; Knight and White, 1989).

In soil analysis, the fraction of carbon containing charred particles is defined as ‘black carbon’ (BC). Thanks to its high stability over centuries, the estimation and identification of BC in soil allowed researchers to define the anthropogenic impact in the pedogenesis processes during the time of occupation of a site (He and Zhang, 2009;

Mazurek et al., 2016; Rumpel et al., 2006). However, a general definition and accepted terminology for BC is still missing (Schmidt and Noack, 2000). Moreover, depending on the methods employed to estimate its content, BC can also include soot and ashes (He and Zhang, 2009), or the heaviest portion of SOC (Gustafsson et al., 2001; Schmidt and Noack, 2000).

Because BC comprises a continuum of compounds, the analytical techniques mainly used are based on chemical treatment of the soil samples for the digestion/degradation of inorganic and organic carbon, coupled with thermal analysis to estimate the BC content (Rumpel et al., 2006; Schmidt and Noack, 2000). As to the analytical conditions, a thermal limit does not exist because of the heterogeneity of the compounds, and this aspect accounts for the lack of standardized analytical procedures for BC determination (most of the time in the thermal analysis the thermal cut-off from BC and organic carbon must be set empirically) (Gustafsson et al., 2001; Schmidt and Noack, 2000).

In order to avoid misinterpretations and erroneous estimates of the BC content, in this research the quantification of charred particles was made using the protocol proposed by Whitlock and Larsen (2002). It is based on a more palynological method of analysis, with no indirect BC quantification. Therefore, the analysis is based on the isolation and quantification of charcoal particles mediated by microscopic observation (for methodological details see paragraph 2.2).

Although this approach is more time consuming and perhaps less efficient than the chemical-thermal one, it is less destructive, and the charcoal fragments can be used for further analysis such as radiocarbon dating or spectroscopic analysis.

Charcoal particles have been widely used for the reconstruction and persistence of past wildfire (Argiriadis et al., 2018; Conedera et al., 2009; McWethy et al., 2020, 2010; Miller et al., 2017, among others) and the identification of the available wood species around a site through the recognition of taxa. Moreover, the change over time in the species of employed wood can inform about the level of technology, the related offer-demand fluctuations of wooden species, and consequently the development of a town even in ancient times (Deforce, 2017). Significantly less attention has been given to the study of charcoal physical and chemical properties to detect the intensity and persistence of fires (Theurer et al., 2021).

Additional information on charcoal structures can be obtained by their analysis through spectroscopic techniques. Raman spectroscopy, in particular, is a non-destructive

technique to determine the thermal maturity of carbonaceous material (Mauquoy et al., 2020) and therefore the Raman analysis of biomass burning residue can be descriptive of the process that produced them.

The properties and characteristics of charcoal -and carbon materials in general- are related to the physical properties of carbon bonds.

The Raman spectrum of charred particles is formed by two bands (further described in section 2.2.5), and the extracted parameters can be tacked back to structural characteristics typical of specific condition of charcoal formation and ageing.

The large amount of charred particles found in Amazonian DE raised the interest of the researchers in different fields, to describe the burned fragments and to find out their impact on Amazonian soils.

Inoue et al. (2017) compared the parameters from charcoal freshly prepared at different temperatures, and charcoals recovered from soil at different depths in Japan. They built a factorial analysis on the parameters and associated the first factor with the charring conditions observing an under-representation of charcoal from soil samples produced at high temperatures. The authors attributed such evidence to the stability of charcoal formed at various temperatures: even if the charcoal originated at higher temperature is chemically more stable because it contains more graphitic domains which rise the stability of the structure, it is also more prone to disappear into the soil because of its physical properties (Ascough et al., 2010; Cohen-Ofri et al., 2006). In fact, higher temperature charring conditions result in more fragile charcoal due to a higher porosity and presence of micro-cracks in wooden cell walls. Furthermore, more graphene-like domains mean also more graphene layers held together by Van der Waals interactions (Chia et al., 2012; Gundale and DeLuca, 2006; Inoue et al., 2017; Keiluweit et al., 2010; Yu et al., 2006). In addition, as demonstrated by Rumpel et al., (2015), charcoal produced at different temperatures can be separated by water flow. As a result, there can be a differential charcoal removal from soil by water erosion.

Charcoal from Amazonian DE was studied by Ribeiro-Soares et al. (2013) through Raman spectroscopy to describe the structure of that particular finding. They were able to determine the crystallite size of natural and laboratory made charred particles and highlight that the differences between the two types of samples were mostly related to the dimensions of the in-plane crystallites. This phenomenon is due to the oxidative degradation of Amazonian charcoal in soil over years of burying. Ribeiro-Soares and colleagues attributed the strong fertility of Amazonian DE soil to the presence of such degraded charcoal. In fact, the increased soil fertility would be due to different

processes: the broken bonds released by the oxidation of aromatic structures of charcoal allow the link with nutrients enhancing the cation exchange capacity of the soil; the small dimension (2-8 nm) of the crystallites permits the diffusion of elements such as Ca, P, and O from the carbonaceous structure to the soil; the aromatic condensed structures possibly act against microbial degradation and leaching into the soil.

Another research focused on the characterization of the charred particles from Amazonian DE was conducted by Jorio et al., (2012). They used a microscopic approach (SEM) to investigate the external surface of carbon grains mapping the elements using the EDX probe. Ca, P, Fe, Al, and Si were commonly detected elements absorbed by the porous surfaces of the samples. Raman spectra were acquired from both the surface and the core of charcoal particles, and the size of the lattice (L_a) was calculated as well. They identified the presence of smaller crystallites on the surface than in the core of the samples and, associating this aspect to the SEM-EDX analysis, illustrated the possibility that this difference is responsible for the retention of the ions on the surface. Comparing the Amazonian DE charred particles to other particles, both natural (turf) and synthesized, they identified these characteristics as unique to the Amazonian soils. Pandey et al., (2020) corroborated the hypothesis of the nutrient retention from charred particles in Amazonian DE highlighting the presence of more sp^3 hybridized carbons (structural defects) in Amazonian DE samples with respect to black carbon from surrounding soils. In addition to the cited researches considering Amazonian DE, Raman spectroscopy has been also employed to study charcoal particles coming from several European sites not strictly related to the DE contexts. De Sousa et al., (2020) studied the Raman spectra and X-ray diffractograms of a series of charcoal particles from archaeologic Anthrosol from different epochs (from ~13,000 years BP to ~500 years BP). Their analysis involved the preparation of freshly made charred particles at controlled heating conditions and statistical analysis. The authors investigated the structural disorder of particles of different sizes and acquired the spectra from different points through a cross-section analysis: smaller particles and surfaces are the most defected, similarly to the results from Jorio et al., (2012). From the comparison between charcoal particles dated back to different periods, they reconstructed the alteration path: the first 2,000 years were affected by an oxidative process that mainly affected the amorphous parts of the fragments (with a relative increase of the nanographitic structures), then the more stable and ordered structure was slowly altered taking an amorphization pathway during the next 9,000 years.

Cohen-Ofri et al., (2006), evidenced an aging process similar to the second step described by De Sousa et al., (2020). They studied the Raman spectra of archaeological charcoal particles dated back to ~3000 years BP, ~ 9000 years BP, and fragments spanning between ~21,000 and 40,000 years BP and recognized a way of degradation of the graphite phase that turns into humic substances.

Theurer et al., (2021) conducted deep research on the Raman parameters (such as the type, positions, and sizes of the characteristic bands) changing under different thermal, precursor and maturity conditions. The authors found significant discrepancies in the behavior of Raman parameters with the variation in the cellulose, hemicellulose, and lignin content of the original plant material. Similar conflicting results were highlighted for the determination of the charcoal maturity: the ratio between amorphous/crystalline fraction in the original materials affects the Raman parameters in complex and still partially unclear ways. From the comparison with previous studies, the work from Theurer et al., (2021) highlighted the difficulties in comparing their results with data from the literature because of the variety of deconvolution methods applied to the Raman spectra. This aspect is not negligible: Raman spectra must be manipulated to extract the parameters needed for comparisons or statistical analysis. The number of curves and their predetermined parameters play a considerable role in the subsequent consideration of any research.

1.2.2 Biomass burning biomarkers: polycyclic aromatic hydrocarbons and anhydrous sugars

During the process of biomass burning several reactions are involved. The incomplete combustion of organic matter produces a high amount of CO_2 and reactive species that react producing highly specific and stable proxies.

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous in the environment and can derive from different processes and sources such as unburned fossil fuel, diagenesis, and mostly from the incomplete combustion of biomass (Lima et al., 2005).

PAHs are constituted by two or more condensed aromatic hydrocarbons, sometimes containing aliphatic branches. Their chemical nature makes them highly stable, and consequently not sensitive to diagenetic alteration. During a burning event, in conditions of oxygen deficiency, organic compounds are fragmented into small and highly unstable free radicals that can react with each other following different paths,

forming the first aromatic ring. This structure develops by reacting with other radicals leading to the subsequent condensation and growth of more stable and complex multi-ring molecules (Lima et al., 2005; Ravindra et al., 2008). Figure 1.1 reports a scheme of the formation and growth of PAHs.

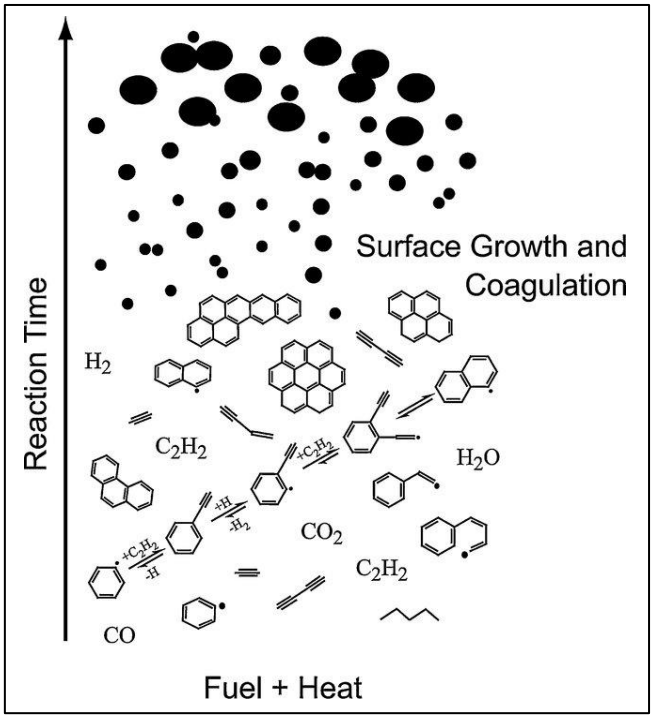

Figure 1.1. Schematic representation of the PAHs formation during burning process, from Lima et al., (2005)

The effects of different reaction conditions (i.e., temperature, time, sample size) and the yields of different PAH congeners were studied by McGrath et al., (2003), demonstrating that for temperatures lower than 300 °C, the combustion produces non-detectable amounts of PAHs. From 400 °C, PAHs with 2, 3, and even some congeners with 4 condensed rings, are produced. The highest amount of 4-rings and some 5-rings congeners is found starting from 500 °C. The heaviest congeners were produced at T ≥ 500 °C together with fluorene and phenanthrene, the most abundant PAHs recognized at 600-650 °C. Figure 1.2 reports the structure of the PAH increasing the number of condensed rings. Different congeners are also differently distributed in the environment: low-weight PAHs are preferentially found in the gaseous phase, while multi-rings congeners are found adsorbed onto solid particles and are defined as “low mobility PAH”. The lighter PAHs can, in turn, undergo worldwide mobility and are preferentially accumulated in the polar regions, being redeposited by means of dry (or

wet deposition (preferential for low weight congeners) (Ravindra et al., 2008). Being more prone to the transport within environment compartments and to photochemical degradation, lighter PAHs are less indicative of fires and burning events that occurred in the past (Argiriadis et al., 2018).

Diagnostic PAH ratios have been proposed to distinguish the source of these organic tracers. Hence, it is possible to distinguish fossil fuel combustion from wood combustion and between hardwood and softwood biomass burning. The values and references of the ratios most frequently used and reported in literature are summarized in the Table 1.1.

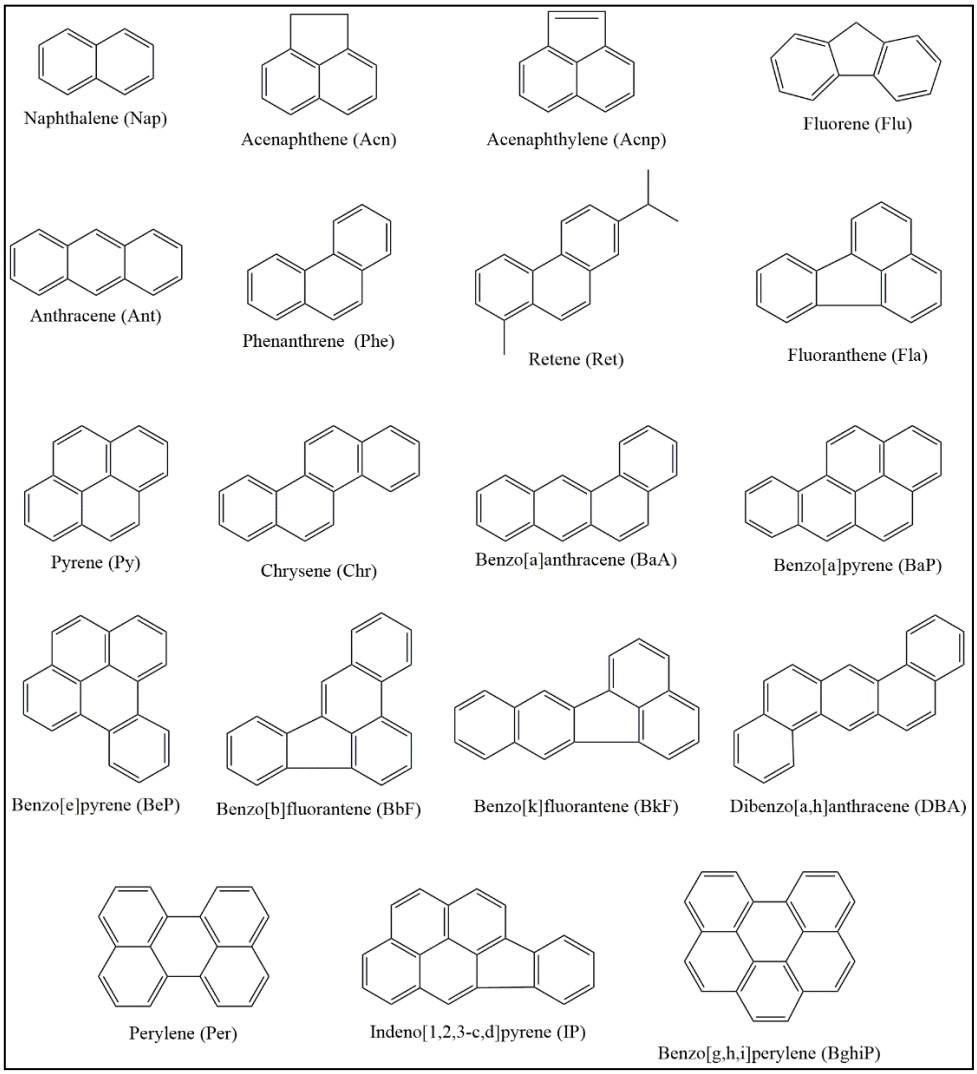

Figure 1.2. Structures, names, and abbreviations of the PAHs determined and discussed in this thesis.

Table 1.1. Ratios of PAH for the identification of the main input.

Ratio	Source	Threshold	Reference
Ant/(Ant + Phe)	Pyrogenic source	>0.1	Yunker et al. (2002)
(IP/(IP + BghiP))	Grasses, wood, coal combustion	>0.5	Yunker et al. (2002)
	Fossil fuel combustion	0.20<x<0.50	
	Petrogenic source	<0.2	
	Cereal straw combustion	0.52<x<1	Hays et al., 2005
Ret/(Ret + Chr)	Petroleum coal combustion	0.15<x<0.50	Kuo et al., 2011; Yan et al., 2005
	Softwood combustion	>0.8	
Fla/(Fla + Py)	Hardwood combustion	0.50<x<0.55	Wang et al., 2009
	Shrubs	0.54<x<0.60	
	Grasses	0.53<x<0.63	Oros et al., 2006
	Cereal straw	0.50<x<0.53	Hays et al., 2005

Anhydrous monosaccharides are the by-products of the incomplete combustion of cellulose. Cellulose, hemicellulose, and lignin constitute the plant cell walls. Levoglucosan (LEVO), the progenitor of anhydrous sugars, is the direct product of thermal degradation of cellulose at temperature higher than 300 °C (Simoneit, 2002; Simoneit et al., 1999). The production of LEVO occurs throughout (i) the depolymerization of cellulose by trans-glycosylation, (ii) the formation of the pyranose intermediate *via* fission and disproportionation bring to the synthesis of levoglucosan and smaller amounts of the two isomers mannosan (MANNO) and galactosan (GALA) (Elias et al., 2001; Shafizadeh, 1984; Simoneit, 2002); the three isomers are reported in Figure 1.3.

At temperatures lower than 300 °C, biomass burning involves depolymerization, water removal and oxidation that lead to char formation. Other polymeric carbohydrates, such as starch, can produce levoglucosan in condition of high temperature degradation, but the temperatures of baking, cooking or toasting temperatures are not enough to depolymerize sugars and produce levoglucosan (Simoneit, 2002). The specificity of LEVO has been demonstrated because it cannot be produced by other cellulose degradation processes or by combustion of fossil fuels (Oros et al., 2006; Simoneit et al., 1999).

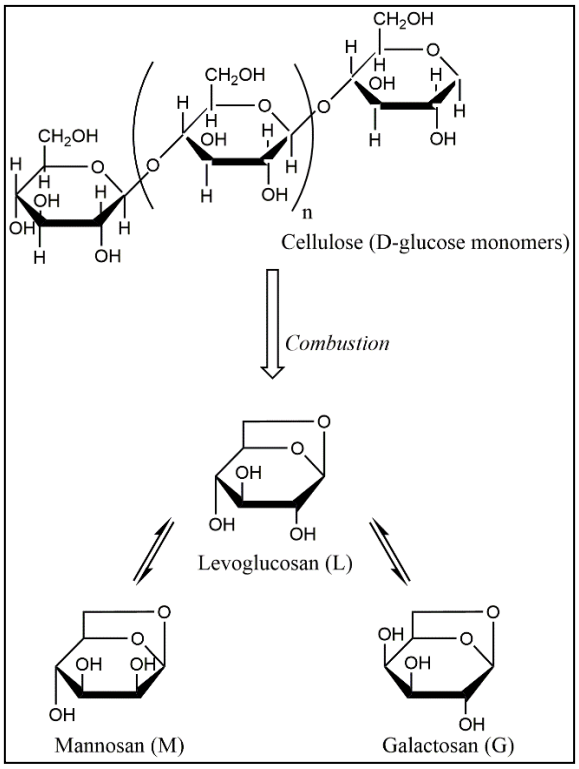

Figure 1.3. Scheme of the combustion of cellulose and structure of the monomers: levoglucosan, mannosan, and galactosan.

The isomers MANNO and GALA have been proposed as deriving preferentially from the thermal degradation of hemicellulose rather than cellulose, hence, the ratios between LEVO, MANNO and GALA have been proposed to be diagnostic of the combustion of gymnosperm and angiosperms (Simoneit, 2002). In fact, several studies report the use of different ratios to distinguish the fuel sources. These ratios represent the amount of hemicelluloses with respect to cellulose in softwood and hardwood. The latter is represented by different values for the ratio LEVO/MANNO: ~5 for softwood, ~10-30 for hardwood and grasses (Alves et al., 2010; Caseiro et al., 2009; Fabbri et al., 2009; Louchouarn et al., 2009; Schmidl et al., 2008; Ward et al., 2006). The ratio LEVO/(MANNO+GALA) has been proposed by Fabbri et al., (2009) to discriminate lignite from other biomasses since no traces of GALA have been detected in lignite burning samples.

The possibilities offered by both groups of proxies (PAHs and anhydrous sugars) raised the interest of the researchers for their possible use in paleoenvironmental reconstructions. Indeed, their concordance or discordance can help elucidating past fire events and even human presence. Many studies used PAHs and sugars to reconstruct

regional and global fire regimes in the past (Battistel et al., 2018; Callegaro et al., 2018; McWethy et al., 2020; Miller et al., 2017; Musa Bandowe et al., 2014; Ruan et al., 2020; Schreuder et al., 2018), first human settlement and the early steps of civilization in isolated or climatic unfavorable areas of the planet (Argiriadis et al., 2018; D'Anjou et al., 2012; Schüpbach et al., 2015).

1.2.3 Linear alkanes

Aliphatic hydrocarbons are lipid molecules biosynthesized by plants (both terrestrial and aquatic), algae, bacteria, and plankton. In terrestrial plants, these molecules mainly constitute the hydrophobic barrier of leaves (Canuel and Martens, 1993; Ficken et al., 2000). Different lengths of these molecules have been associated with different vegetal (or bacterial) organisms. The interest of the environmental researchers in this group of proxies lies in their versatility to describe the typology of vegetation and the resulting environmental and climatic conditions of an area. The different *n*-alkane pattern can provide the relative proportion of grasses, trees, and aquatic plants. Moreover, the detection of different *n*-alkanes into natural continuous archives (i.e., soil and sediment core) can reveal a regional to local sequence of dry and wet periods (Callegaro et al., 2018; D'Anjou et al., 2012; Fang et al., 2014; Silliman et al., 1996; Silva et al., 2012; Zech et al., 2010 among others). Table 1.2 reports the different sources attributed to various *n*-alkanes pattern.

Table 1.2. Vegetal and bacterial source associated to *n*-alkanes length.

***n*-alkane**	**Source**	**Reference**
$< C_{20}$ (even number)	Bacteria	Cranwell et al., 1987; Fang et al., 2014; Meyers, 2003; Pisani et al., 2013
$< C_{20}$ (odd number)	Plankton	
$C_{20} - C_{25}$	Submergent and floating aquatic plants	Cranwell et al., 1987; Ficken et al., 2000
C_{27} and C_{29}	Trees and shrubs	Fang et al., 2014; Ficken et al., 2000; Killops and Killops, 2005; Oros et al., 2002; Silliman et al., 1996; Silva et al., 2012; Zech et al., 2010
C_{31} and C_{33}	Grasses and herbs	Killops and Killops, 2005; Zech et al., 2010

In addition to the net abundance of *n*-alkanes in the sample, different ratios for source identification have been proposed to distinguish sources and development processes.

Average chain length (ACL) for example, is the expression of the weighted mean abundance of the alkanes in a sample. It is calculated as reported in the equation (1):

$$\sum(C_i \times i)/\sum C_i \quad (1)$$

where C_i is the concentration of the *n*-alkanes with *i* number of carbon atoms (Derrien et al., 2017; Fang et al., 2014). ACL is used to quickly identify the prevalent source of linear alkanes.

Another parameter frequently used to determine the type of ageing of the organic matter is the odd-even predominance (OEP). OEP indicates the abundance of *n*-alkanes with an odd number of carbon atoms with respect to the congeners with an even number of carbon atoms. The latter, indeed, represent the result of microorganism re-working of the alkanes. OEP is defined by the following equation (2):

$$OEP = ([C_{27}] + [C_{29}] + [C_{31}] + [C_{33}])/([C_{26}] + [C_{28}] + [C_{30}] + [C_{32}]) \quad (2)$$

Generally, a higher value of the OEP is related to a low rate of degradation of the organic matter by microorganism, whereas low values of OEP are associated to a strong microorganism re-working (Zech et al., 2013, 2010).

1.2.4 Miliacin

Millet was one of the first species domesticated by humans starting in China between 8,000 and 10,000 years ago (Lu et al., 2009). It has been known since the beginning of human history thanks to the high resistance characteristics of this crop. The importance of millet cultivation consists in the fact that it has the lowest water requirements among all the other grain crops (Baltensperger, 2002), and it can develop in very poor soil and with a very fast growth rate (Lu et al., 2009; Nesbitt and Summers, 2022). The presence of traces of this cereal has been recorded mostly in Asia and Europe throughout all human history since early settlement (Castelletti et al., 2001; Hunt et al., 2011, 2008; Zhao, 2011). Its relevance is testified by the fact that during the Middle Ages in Europe, millet was tax-free to guarantee access to this resource to the people from the working classes (Ganzarolli et al., 2018).

Miliacin is a pentacyclic triterpene (reported in figure 1.4) produced almost uniquely by millet, as an antibacterial and cytotoxic agent (Bossard et al., 2013), indeed the abundance of miliacin in millet can explain why this plant is known since antiquity and maintained up to present (Motuzaite-Matuzeviciute et al., 2013).

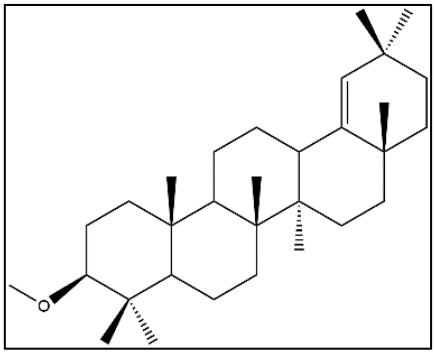

Figure 1.4. Structure of miliacin.

Since its discovery, miliacin has been used to trace the presence of agricultural practices (Bossard et al., 2013; Jacob et al., 2009, 2008), and consequently its fluctuation in natural archives can be a direct sign of the change in human activities. Some past studies quantified the presence of miliacin in soil and sediment samples starting from the Bronze Age.

Bossard et al., (2013) focused on the relative presence of this molecule in different parts of *Panicum miliaceum* (broomcorn millet) to detect where miliacin is mainly stored inside the plant. They observed that this molecule is present in seeds, leaves and stems, and almost absent in hulls and roots. As reported by Jacob et al., (2008), during the transition between the Bronze and Iron ages miliacin concentration increased in the alpine area of France, highlighting a wide diffusion of the cultivation of millet in central Europe. On the other hand, during the Middle Ages its concentration decreased slowly. This because between the Iron and Middle Ages, surprisingly, Romans used millet just as a minor crop: this cereal is rarely detected during the Roman period (Murphy, 2016).

1.2.5 Fecal sterols and stanols

The best way to detect the presence of a living being, both in the past and in the present, is by researching its products. Fecal proxies are some of the most used analytical

biomarkers to detect the presence of fecal pollution from humans and animals in environmental matrices.

Fecal sterols and stanols are widely used molecules for the determination of sewage pollution in water since the first studies by Leeming and colleagues around the '90s (Leeming et al., 1996, 1994).

Sterols are polycyclic organic congeners derived from sterol, characterized by a condensed 4-ring structure and a hydroxyl group. Sterols are defined as amphipathic lipids as they are characterized by a lipophilic structure, but they also have a polar group. They structurally derive from 1,2-cyclopentanophenanthrene and the IUPAC nomenclature refers to the rings designated as A, B, C, and D. The carbon numbers used in this work are the ones proposed by Goad and Akihisa, (1997). Figure 1.5 reports the base structure of 1,2-cyclopentanophenantrene and the nomenclature of the cholesterol molecule. All the substituents can be axially oriented below (α) or above (β) the plane of the rings. The presence of double bonds is typical of some common sterols and stanols, and the capital Greek letter delta (Δ) is normally used to indicate the presence of molecules with one or more unsaturated bonds, the number after Δ defines the position of the double bond (Goad and Akihisa, 1997).

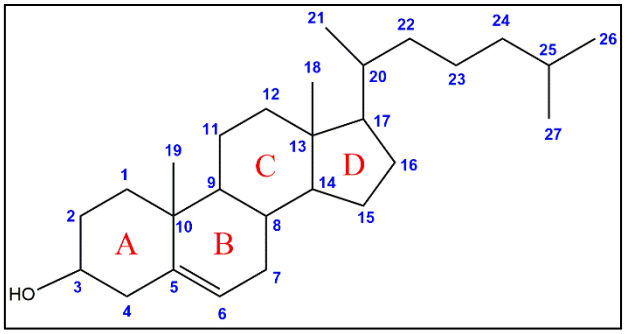

Figure 1.5. Structure of cholesterol molecule with the carbon numbers following Goat and Akihisa (1997).

In eukaryotes, sterols are synthesized by acetyl-coenzyme A from squalene. They play an essential role in the structure of the cell membrane and in cellular metabolism as precursors of steroid hormones. Cholesterol is the most important sterol for human beings as it is responsible for the synthesis of hormones such as corticoids (C_{21}), mineralocorticoids, progestogens (C_{21}), androgens (C_{19}), and estrogens (C_{18}), but also vitamin D and bile acids (Goad and Akihisa, 1997; Juste and Gérard, 2021).

In the intestine of mammals, reduction reactions of the sterols occur, mediated by enteric bacteria. Δ^5-sterols lose an unsaturation and become preferentially 5β-stanols.

5α-stanols can be formed in conditions of anaerobiosis (such as deep in soil or sediments) starting from the β isomers (Bull et al., 2002; Hatcher and McGillivary, 1979; Murtaugh and Bunch, 1967). The mechanism of reduction of Δ-sterols involves some intermediates such as cholest-5-en-3-one and 5β-cholestan-3-one and is reported in Figure 1.6 (Björkhem and Gustafsson, 1971; Bull et al., 2002; Dewei et al., 1996; Eyssen et al., 1973). Coprostanol is the most abundant 5β-stanol with respect to the total sterol content in human feces (Bull et al., 2002; Leeming et al., 1994). In feces of other mammals (such as sheep and cows) the most abundant sterols and stanols derive from an herbivorous diet. For these animals, the Δ^5-sterols can be campesterol or sitosterol (instead of cholesterol) and the derived 5β-stanols would be preferentially 5β-campestanol and 5β-stigmastanol (Bull et al., 2002; Leeming et al., 1996).

Due to their chemical structure, sterols and stanols are low hydrophilic compounds that preferably associate with sedimentary particulate matter which, in turn, allow for a longer preservation of these proxies even in aqueous environments (Bull et al., 2002). The specificity of these compounds for the determination of different fecal inputs and their persistence in the environment raised the interest of the researchers for their application in pollution sciences and paleoenvironmental reconstructions.

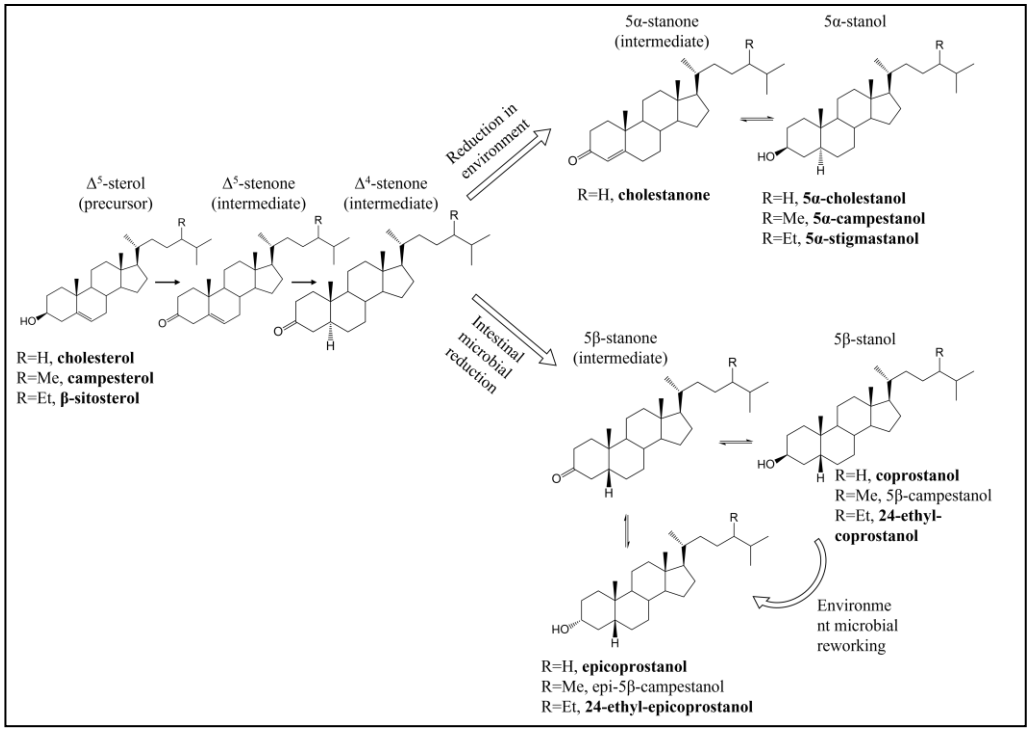

Figure 1.6. Schematic path of transformation of sterols in 5α and 5β stanols in environment and animal gut. Modified from Bull et al. 2002. The names in bold are the compounds investigated in the present work.

As already mentioned above, Leeming et al. (1996) proposed the use of specific sterols and stanols to detect human fecal pollution. Thus, coprostanol represents ~60% of the total content of sterols and stanols and has been proposed as a powerful proxy for human pollution quantification in wastewater (Carreira et al., 2004; Chan et al., 1998; Grimalt et al., 1990; Leeming et al., 1996; Reeves and Patton, 2005; Wu et al., 2009). The non-ubiquity of this molecule was questioned by some other researchers because of the presence of coprostanol in the feces of other mammals such as marine mammals, pigs, cats, and dogs (Jardé et al., 2007b; Leeming et al., 1996; Tyagi et al., 2007; Venkatesan and Santiago, 1989). Therefore, the studies focused on the identification of specific ratios of sterols and stanols able to describe the 'sterol fingerprint' of sources instead of a unique compound (Derrien et al., 2017). Indeed, the distribution of sterols and stanols in animals depends on different factors, among which (i) the diet (omnivore, herbivore), (ii) the metabolism of the animal and its ability to biosynthesize sterols, and (iii) the intestine microbiome capability to reduce sterols into stanols and related compounds (Bull et al., 2002; Leeming et al., 1996). Table 1.3 reports the names, abbreviation, and classification of the fecal sterols and stanols that will be discussed in the following chapters.

Table 1.3. Names, abbreviation, and classification of the fecal sterols and stanols.

Trivial name	Abbreviation	IUPAC name	Classification
Cholesterol	CHL	Cholest-5-en-3β-ol	Δ^5-sterol
Cholestanol	CHN	5α-cholestan-3β-ol	5α-stanol
Coprostanol	COP	5β-cholestan-3β-ol	5β-stanol
Epi-coprostanol	ECOP	5β-cholestan-3α-ol	5β-stanol
Campesterol	CAMPE	Campest-5-en-3β-ol	Δ^5-sterol
β-sitosterol	β-SITO	Stigmast-5-en-3β-ol	Δ^5-sterol
Stigmastanol	STGN	5α-stigmastan-3β-ol	5α-stanol
Stigmasterol	STGR	(22E)-stigmasta-5,22-dien-3β-ol	Δ^5-sterol
24-ethyl-coprostanol	24-COP	24-ethyl-5β-cholestan-3β-ol	5β-stanol
24-ethyl-epi-coprostanol	24-ECOP	24-ethyl-5β-cholestan-3α-ol	5β-stanol
Cholestanone	CHONE	5α-cholestan-3-one	stanone

Derrien et al (2017) define the combination of the three abovementioned elements as a 'steroid fingerprint', illustrating differences among some of the main species: herbivores' feces are rich in sterols and stanols with a structure composed of 29 carbon atoms (β-sitosterol, 24-ethyl-coprostanol, stigmastanol), whereas those of omnivores

are characterized by 27-carbon congeners such as cholesterol, coprostanol, and epicoprostanol (Derrien et al., 2011; Evershed et al., 1997; Jardé et al., 2007a; Leeming et al., 1996; Shah et al., 2007; Tyagi et al., 2008).

Derrien et al., (2017) proposed an extensive review of the main ratios of sterols and stanols, developed during the last 25 years of research and, in parallel, introduced some new approaches developed since the early 2000s: statistical and chemometric strategies have been improved using fecal proxies to point out the different pollution sources (animals, humans).

For a deeper investigation of the ratios and their meaning, please refer to the section 5.2.

Goals and objectives

The dark earth layers that will be discussed in this thesis were excavated in the town of Verona, located in north-eastern Italy. The sites of interest are three: (1) Verona city center-vicolo San Pietro in Monastero (SPM); (2) Verona city center-via Pigna (VP) and (3) the historical village Castelar di Leppia (Figure 1.7.) The archaeological excavations carried out in the old town in Verona exposed DE layers dating back to the Middle Ages, while the layer sampled in Castelar di Leppia dates back to the Bronze Age. VP and SPM sites were chosen because representatives of a context of reuse of past roman structures during the medieval period, whereas CDL represents a DE of a different epoch, but referred to the same area and similar in terms of geological setting conditions.

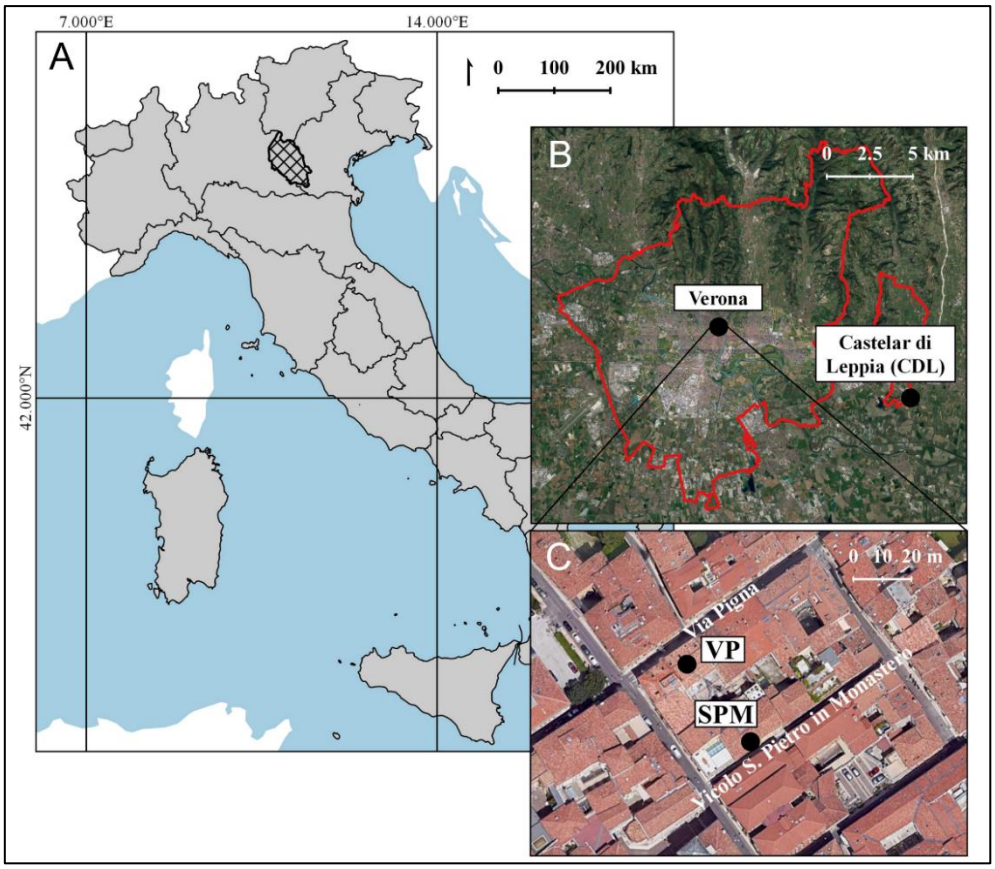

Figure 1.7. Location of the studied sites. A. gridded portion of Italy represents the province of Verona; B. Verona city center and Castelar di Leppia (CDL) locations from GPS image ESRI; C. GPS image of the via Pigna (VP) and vicolo San Pietro in Monastero (SPM) site location.

The aims of this project are (i) to provide a detailed description of these DE soils and (ii) to propose a tentative mechanism for their formation, involving both natural and anthropogenic factors in order to reconstruct the activities involved in the establishment and development of these layers This will be achieved by determining the proxies described so far and consequently identifying signature features characteristic of each site.

As widely described in the previous sections, the study of cultural layers requires a multi-analytical approach due to their intrinsic complexity. Therefore, a micromorphological soil characterization, with the analysis of inclusions and trace elements, will be carried out. In addition, the literature suggests that the most prominent hypotheses of the DE layers formations lie on (i) the strong input of biomass burning, (ii) the presence of cultivated areas, and (iii) the occurrence of manure and/or domestic animals in the studied site. To evaluate all these possible contributions to the proxies, an integrated approach will be followed. The input given by burning residues will be evaluated through the determination and quantification of charcoal; in parallel, an integrated analysis of molecular proxies derived from the combustion of biomass (PAHs, anhydrous sugars) will be performed. The vegetation coverage will be determined by the analysis of linear alkanes and miliacin, while the presence of animals and human excrements will be determined through the analysis of fecal sterols and stanols.

2 Materials and Methods

Parts of this section were previously published in: Bortolini, M.*, Agnoletto, F. C., Argiriadis, E., Nicosia, C., McWethy, D. B., Devos, Y., Stortini, A. M., Baldan, M., Roman, M., Vendrame, T, Scaggiante, R, Bruno, B, Pojana, G., Battistel, D. Insight into the carbonaceous fraction of three cultural layers of different age from the area of Verona (NE Italy), CATENA, 217 (2022), (DOI: 10.1016/j.catena.2022.106453). The work was open access published and under the CC BY license.

The undersigned contributed to the paper for the investigation, formal analysis, writing, and supervision.

2.1 Study sites

The area of the province of Verona consists of a mountain region in the north (Baldo Mount and Lessini mountains) and the plane formed mainly by deposits of the Adige River. The plane formation has a quote that varies from 20 to 200 m, mainly formed by calcareous gravel dated back to the Last Glacial (De Zanche et al., 1977; Zorzin, 2021). Rainfall is attested around 700-1500 mm mainly in spring and autumn, annual mean temperature is around 12-13 °C (data from web repository of ARPA Veneto).

Undisturbed blocks for soil micromorphology were collected from the exposed stratigraphic profiles: Figure 2.1 shows the location of the collected blocks to be analyzed in thin section. The blocks were successively aliquoted to provide samples for the other analysis. Specifically, two samples were collected from each section: one from the top and one from the bottom of the block. The amount of material for each sample was calibrated to ensure the maximum integration between all analytical techniques.

The first excavation site is in vicolo San Pietro in Monastero (SPM) (45° 26′ 42.416″ N, 10° 59′ 48.098″ E) and contains two parallel series of samples (SPM I and II).

At Via Pigna site (VP) (45° 26′ 43.241″ N, 10° 59′ 47.898″ E), the sampled sequences VP I and VP II intercepted the same portion of stratigraphy. A mortar floor was in fact the marker layer allowing to correlate the two sequences. Above this floor a 20–25 cm layer of very charcoal-rich DE was sampled.

At the site Castelar di Leppia (CLD) (45° 24′ 29.312″ N, 11° 8′ 20.891″ E) a ca. 40 cm-thick, dark grey cultural layer was sampled. This derives from the welding of two anthropogenic layers, that are separated by the wide earthen rampart that surrounded the site during the Bronze Age.

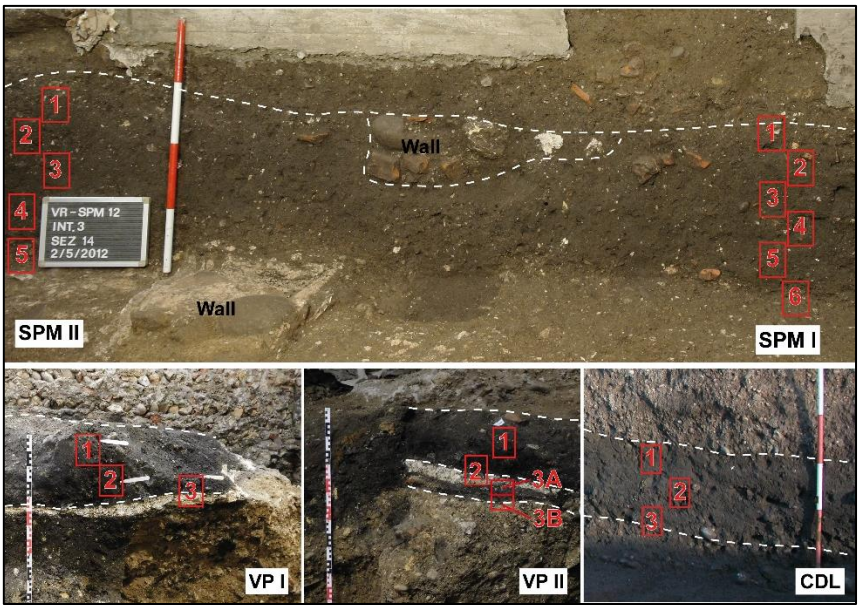

Figure 2.1. Profiles of sampling of the blocks for the preparation of the thin sections. From the blocks top, bottom or bulk aliquots were collected to perform the analysis.

2.2 Soil characterization

The materials and applicative methods described in this section include the microscopic procedures and elemental analysis followed for the characterization of the samples, comprising the quantification and Raman analysis of the charred fraction present as inclusion in the soil samples.

2.2.1 Soil Micromorphology

Thin sections (60 × 80 mm) were prepared from air-dried undisturbed blocks following the laboratory procedures of Beckmann (1997) and described using the terminology of Stoops (2003). Thin sections were studied in plane polarized (PPL) and cross polarized (XPL) transmitted light with magnifications ranging from 12.5x to 400x. A total of twenty-one thin sections were produced and observed: six from the profile SPM I, five from the profile SPM II, three and four from the profiles VP I and VP II respectively, and three from CDL sampling site. The arrangement and size of black particles in the

three sets of samples was determined by observing the groundmass at 100x - 200x magnification and with the use of the sub-stage condenser.

2.2.2 Scanning electron microscopy (SEM)

This analysis was carried out by isolating the particles of interest and sticking them manually to a carbon conductive tape. SEM and energy dispersive spectroscopy (EDS) analysis were performed using a TM3000 tabletop scanning electron microscope (Hitachi, Tokyo, Japan) coupled to an X-ray microanalysis system SwiftED3000 (Oxford Instruments Inc., Bognor Regis, UK). Conditions for recording the EDS spectra were as follows: acquisition time 30.0 s, process time 5 s, accelerating voltage 15 kV.

2.2.3 Trace Elements

The elemental analysis of Fe, Pb, Cu, Mn, Sr, Ba and P was carried out in 4 samples from each SPM and VP profiles, and in 1 sample from CDL. Approximately 0.25 g of homogenized soil samples were microwave digested in 3:1 v/v HCl:HNO_3, adapting the method described in Argiriadis et al. (2021). Trace elements were determined using an Agilent (Palo Alto, CA, USA) MP-4210 MP-AES instrument fitted with a double-pass cyclonic spray chamber and a OneNeb Series 2 nebulizer. The selected wavelengths and MP-AES operating conditions are reported in Table 2.1.

Table 2.1. MP-AES operating conditions *.

Element	Wavelenght (nm)	Nebulizer flow (mL min^{-1})	Read time (s)
Ba	455.403	0.65	5
Cu	324.754	0.70	1
Fe	371.993	0.95	5
Mn	403.076	1.00	5
P	214.915	0.45	5
Pb	363.957	0.75	3
Sr	460.733	1.00	1

* Common conditions: replicates 3; pump speed 15 rpm; uptake time 15 s; stabilization time 30 s.

Quantification was carried out by 8-points external calibration standards prepared from a mixture of the multielemental ICM-103 solution (Ultrascientific, BO, Italy), and monoelemental S and P mother solution. Yttrium was used as internal standard. Results

are reported as enrichment factor (EF) calculated using the ratio proposed by Reimann and De Caritat, (2000) and reported in the following equation (3):

$$EF_i = \frac{[i]_{sample} / [Al]_{sample}}{[i]_{ref} / [Al]_{ref}} \quad (3)$$

where [i] is the concentration in the sample of the element and *ref* indicates the use of Earth's crust concentrations (Rudnick and Gao, 2013) as a reference. Aluminium was used for normalization, because it is assumed to exclusively derive from crustal sources in regional soil contexts (Reimann and De Caritat, 2000).

2.2.4 Quantification of the charred particles

The aliquots were dried at room temperature in a desiccator where silica gel was changed every 24 h. When constant weight had been reached, three aliquots -each of ~ 0.5 g- were prepared for the analysis of charred particles. During this phase, the biggest fragments of charcoal (>500 μm) were extracted and preserved for further Raman analysis. The rest of the dried sample was sieved and crushed with a pestle and mortar for further biomarker analysis.

The triplicate aliquots for charred particle count were prepared following Whitlock & Larsen, (2002). Briefly, the soil was soaked in a deflocculant solution prepared with 50 mL of sodium hexametaphosphate 10% (crystalline, +200 mesh, 96%, Merck, St. Louis, MO, USA) and 50 mL of sodium hypochlorite (technical grade, Merck, St. Louis, MO, USA) for 24 h and then gently washed through a series of nested sieves (mesh opening sizes of 1 mm, 500, 250, 125 and 63 μm). The fractions were collected in demineralized water into Falcon® plastic tubes with caps until analysis and stored at room temperature.

To determine the content of residues from combustion, the count of the charred particles was done by microscopic observation of the samples previously disaggregated from the matrix. The samples were transferred into a gridded petri dish and for the counting of the charcoal particles. The observation was performed with a stereomicroscope Olympus SZ61 following bibliographic guidelines (Enache and Cumming, 2006; Kirchgeorg et al., 2014; Mustaphi and Pisaric, 2014). All samples were analyzed in triplicate.

2.2.5 Raman spectroscopy

A total of 29 samples (16 from SPM, 10 from VP and 3 from CDL) among the larger (>1 mm) charcoal fragments were selected for μ-Raman analysis. Samples were mechanically cleaned using scalpels and brushes to expose a fresh surface. Spectra were acquired with an i-Raman 785 s spectrophotometer by BWTek (Newark, DE, USA), coupled to its dedicated optical microscope BAC151B Raman Video Microsampling System by means of optical fibre (1.5 m length) ending in a BAC102 Raman Trigger Probe equipped with a standard 304SS shaft mounting a flat quartz window. The Rayleigh radiation was blocked by a notch filter and the backscattered Raman light was dispersed by a holographic grating on a TE Cooled Linear 2048 pixels CCD Array (cooling temperature: 10 ◦C); the entrance slit width was fixed at 25 μm. The laser excitation source consists of a 785 nm diode whose power can be modulated (1% steps) from 3 to 300 mW to avoid thermal effects on the analyzed sample. Precisely, the spectra were collected with a 20X objective at a power range between 7 and 15 mW, in the 175–3000 cm^{-1} spectral range, with typical integration times of 60 s and a nominal spectral resolution of 4.5 cm^{-1}. In order to improve the signal-to-noise ratio, five accumulation cycles were collected for each spectrum. Flat and clean surface portions were accurately selected by means of the microscope camera and each sample was characterized by analysing 2 to 4 surface points.

All raw Raman spectra were then decomposed to identify the position and intensity of the bands according to methodologies proposed by Inoue et al., (2017) and McDonald-Wharry et al., (2013). A routine method, to allow the comparison of spectra recorded with different instruments, was implemented. For each spectrum the spectral range between 900 and 1800 cm^{-1}, where most of the typical carbon features are located (Dennison and Holtz, 1996; Ferrari and Robertson, 2000; McDonald-Wharry et al., 2013) was selected. A linear baseline correction was employed choosing the minimum between 900 and 1000 cm^{-1} and 1600–1800 cm^{-1} as constant anchor points (Inoue et al., 2017). Based on literature evidence, the curve fitting was performed using Gaussian functions employing eight bands: from lower to higher values of Raman shift: SL, S, DS, D, A1, A2, GG and GL (Figure 2.2) (Inoue et al., 2017; McDonald-Wharry et al., 2013).

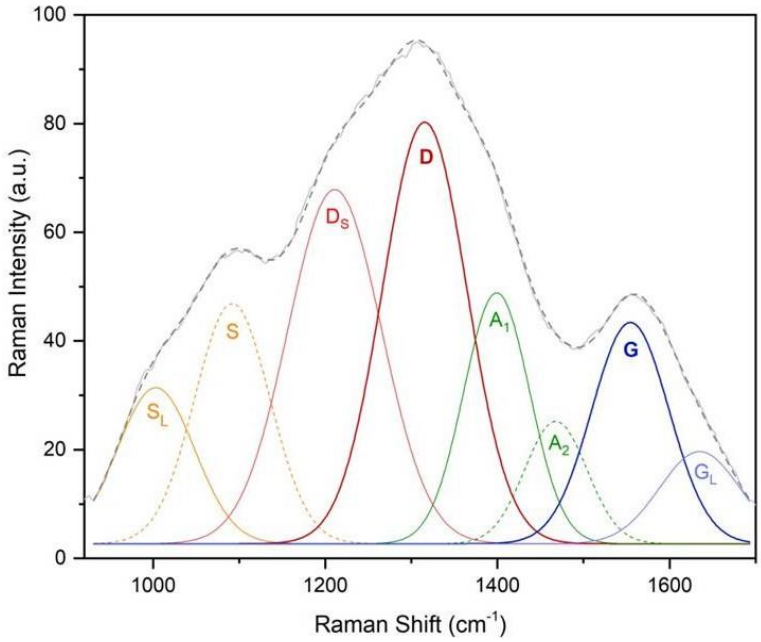

Figure 2.2. typical deconvolution fitting from a charred particle sample of the SPM site. The eight bands used for the deconvolution of the spectra are reported from lower to higher values of Raman shift: S_L, S, D_S, D, A_1, A_2, G_G and G_L).

2.3. Radiocarbon dating

The principal interest of the radiocarbon dating in this context was to define if a chronology could have been defined, and being the excavation profiles of VP and SPM referred to the same stratigraphic horizon the best-preserved profile was identified.

SPM II profile showed the less disturbed sequence during the excavation phases; hence, sequence was possibly the most suitable for the determination of a chronology through the soil depth.

Bulk samples and charred fragments were sent to Beta Analytics Laboratories (Miami, FL, USA) for analysis. They provided the humin fraction extraction from the total soil carbon using acid-alkaline treatment from 5 top aliquots from SPM II soil samples. The procedure consists of first sieving of the soil, then acid treatment is performed to remove carbonates, the solution is thoroughly washed, and followed by an alkali treatment aiming to solubilize humic fraction. The alkali-insoluble fraction is the humin fraction ready to be dated. From the same profile (i.e., SPM II), 3 charcoal fragments were hand-picked for radiocarbon dating. Both humin fraction and charcoal fragments

were analyzed through Accelerated Mass Spectrometry at the Beta Analytics Laboratories (Miami, FL, USA). Conventional Radiocarbon Ages (CRA) were calibrated using BetaCal4.20: HPD method: INTCAL20.

2.4. Organic, inorganic and recalcitrant carbon

The total organic carbon (TOC), recalcitrant organic carbon (ROC) and total inorganic carbon (TIC) were determined through a Soli TOC Cube (Elementar, Langenselbold, Germany). The method employed (DIN 19539:2016–12) consists of a temperature ramp from 150 °C to 900 °C, using oxygen and nitrogen as reactive and carrier gas, respectively. The dry combustion of TOC occurs between 150 and 400 °C, while ROC is combusted between 400 and 600 °C and TIC is later released up to 900 °C.

2.5 Biomarkers

The materials and procedures that concern the molecular proxies, namely biomarkers, are divided following their chemical nature: apolar or polar. The preparation and analysis sections are thus divided according to this classification.

To guarantee minimal contamination of the samples, only dedicated glassware lab equipment was used. Before every use, decontamination rinses were performed using methanol (MeOH), dichloromethane (DCM), and *n*-hexane (HEX) (pesticide grade, Romil Ltd., Cambridge, UK). For Pasteur pipettes and inert powders, decontamination was performed for 4 hours at 400°C in a furnace oven.

2.5.1 Analysis of the apolar fraction

The preparation and analysis of the samples was performed adapting the procedure from Battistel et al., (2017).

The extraction for the analysis of the apolar fraction (PAHs, *n*-alkanes, sterols and stanols, and miliacin) was performed with an accelerated solvent extractor ASE 200 (Dionex Thermo Fisher Scientific, Waltham, MA, USA) equipped with 12 (22 mL volume) stainless steel cells. ~ 5 g of powdered samples were weighted, mixed with: diatomaceous earth (Thermo Fisher Scientific, Waltham, MA, USA) used as filling inert material, and 2 g of anhydrous sodium sulfate (pharmaceutical secondary standard, Merck, St. Louis, MO, USA) preserved in a hermetic bottle with HEX (pesticide grade, Romil Ltd., Cambridge, UK). At the bottom of the cell, a 22 mm diameter glass fibre filter (Thermo Fisher Scientific, Waltham, MA, USA) was placed. The extraction program consisted of two 5 minutes static cycles at 150 °C at 1500 psi. The extraction solvent was a mixture of DCM and MeOH 9:1 (Battistel et al., 2015); both solvents were pesticide grade (Romil Ltd. Cambridge, UK). Before the extraction, the internal standards were added to the samples:

- For PAHs quantification, 50 absolute ng of Acenaphthylene $^{13}C_6$ (98% chemical purity, solution 100µg mL^{-1} in nonane, CIL, Cambridge, UK), Phenantrene $^{13}C_6$ (98% chemical purity, solution 100µg mL^{-1} in nonane, CIL, Cambridge, UK), Benzo(a)Pyrene $^{13}C_4$ (98% chemical purity, solution 100µg mL^{-1} in nonane, CIL, Cambridge, UK).
- Hexatriacontane (98%, Merck, St. Louis, MO, USA) was used for the determination of aliphatic hydrocarbons.
- Quantification of miliacin was performed using 5α-androstane (certified reference material 2 mg mL^{-1}, Merck, St. Louis, MO, USA) as internal standard.
- For quantification of sterols and stanols labelled cholesterol ^{13}C was used (Cholesterol-25,26,27-$^{13}C_3$, Merck, St. Louis, MO, USA)

The liquid extracts were collected in 60 mL glass vials and stored in the fridge (4°C). Then, samples underwent volume reduction until ~0.5 mL using Turbovap® (Caliper Life Science, Perkin-Elmer, Waltham, MA, USA) set at 23 °C.

Clean-up was performed using Pasteur pipettes filled with glass wool and ~2 cm of silica. Prior to the use, filled Pasteur pipettes were rinsed with one aliquot both of

MeOH and DCM. The volume-reduced sample was eluted followed by 10 mL of DCM. Finally, the samples were dried under a gentle flux of nitrogen and stored in the fridge (4 °C) until the analysis.

The analysis of the apolar analytes was performed using an Agilent Technologies 7890A GC system coupled to an Agilent Technologies 5975C inert MSD quadrupole mass selective detector. Injection volume was 1 µl in splitless mode, injector was operated at 280 °C. The capillary column used to separate the analytes was a 60 m Zebron ZB-5plus (60 m length, 0.25mm inner diameter, 0.25 µm film thickness) whose stationary phase is composed of (5%-phenyl)-dimethylpolysiloxane. The electron impact source was set at 70 eV and 230 °C, mass spectrometer was operated to positive mode at 150 °C and the acquisition was acquired using selected ion monitoring (SIM) mode.
In the following paragraphs the specific chromatographic and spectrometric parameters for each class of analytes will be examined in detail.

2.5.2.1 PAHs

Inlet temperature was operated at 300°C and sample was injected in splitless mode and the volume of injection was 2 µL, carrier flow (He) was 1 mL min^{-1}.
Temperature program was the following: initial temperature 70 °C held for 1.5 min, then 10 °C min^{-1} until 150 °C held for 10 minutes, then a second ramp at 3 °C min^{-1} until 300 °C held for 15 minutes. Post run was at 305 °C for 30 minutes.
The quantification was calculated using the internal standard methos: Acenaphthylene $^{13}C_6$ was used to quantify naphthalene, acenaphthylene, acenaphthene, and fluorene; labelled phenanthrene was used to quantify phenanthrene, anthracene, fluoranthene, pyrene, and retene; benzo(a)pyrene $^{13}C_6$ was employed to quantify benzo(a)anthracene, chrysene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(e)pyrene, benzo(a)pyrene, perylene, benzo(ghi)perylene, dibenzo(A,H)anthracene, indeno(1,2,3-c,d)pyrene. All the PAHs analyzed with the chromatographic and SIM parameters are reported in Table 2.2.

Table 2.2. PAHs analytes, retention times and target ion for each congener; IS was used to quantify the species above.

Analyte	Retention time (min)	Target ion (m/z)	MW (g mol^{-1})
Naphthalene	12.3	128	128
Acenaphthylene	22.6	152	152
Acenaphthene	22.6	154	154
Fluorene	27.2	166	166
Acenaphthylene $^{13}C_6$ (internal standard)	21.1	158	158
Phenanthrene	35.7	178	
Anthracene	36.1	178	
Fluoranthene	45.7	202	
Pyrene	47.4	202	
Retene	50.3	234	
Phenantrene $^{13}C_6$ (internal standard)	35.7	184	
Benzo(a)Anthracene	57.2	228	
Chrysene	57.4	228	
Benzo(b)Fluoranthene	65.1	252	
Benzo(k)Fluoranthene	65.2	252	252
Benzo(e)Pyrene	66.9	252	
Benzo(a)Pyrene	67.2	252	
Perylene	67.7	252	
Benzo(ghi)Perylene	74.7	276	
Dibenzo(A,H)Anthracene	75	278	
Indeno(1,2,3-c,d)Pyrene	76.6	276	276
Benzo(a)Pyrene $^{13}C_4$ (internal standard)	67.2	256	

2.5.2.2 *n*-Alkanes

For the determination of aliphatic hydrocarbons, the injection volume was 1 µL, in splitless mode, and the temperature of the inlet was set to 300 °C. The oven program was configured with a static at the temperature of 50 °C held for 5 minutes, followed by a ramp at 18 °C min^{-1} until 315 °C, followed by a static of 16 minutes. Post run cycle was set at 315 °C for 15 minutes. Helium flow was set at 1.5 mL min^{-1}.

The *n*-alkanes of interest are the linear hydrocarbons from the *n*-decane ($C_{10}H_{22}$) to the *n*-pentatriacontane ($C_{35}H_{72}$). All the congeners analyzed are reported in Table 2.3.

The concentration of each congener was calculated employing the area and concentration of the internal standard (hexatriacontane) and corrected for the instrumental response factor. All the analytes were integrated using the target ion with m/z 71, which correspond to the loss of five methyl groups from *n*-decane, forming the molecular ion of *n*-pentane. This ion produces an intense line in the mass spectrum of

all the linear alkanes species and it is relatively stable if compared to the other ion fractions.

Table 2.3. *n*-alkanes retention times and molecular weight, SIM target ion is m/z 71.

Analyte	Formula	Retention time (min)	MW
Decane	$C_{10}H_{22}$	10.9	142
Undecane	$C_{11}H_{24}$	12.0	156
Dodecane	$C_{12}H_{26}$	13.0	170
Tridecane	$C_{13}H_{28}$	13.8	184
Tetradecane	$C_{14}H_{30}$	14.6	198
Pentadecane	$C_{15}H_{32}$	15.3	212
Esadecane	$C_{16}H_{34}$	16.0	226
Eptadecane	$C_{17}H_{36}$	16.6	240
Ottadecane	$C_{18}H_{38}$	17.2	254
Nonadecane	$C_{19}H_{40}$	17.8	268
Eicosane	$C_{20}H_{42}$	18.3	282
Eneicosane	$C_{21}H_{44}$	18.8	296
Docosane	$C_{22}H_{46}$	19.3	310
Tricosane	$C_{23}H_{48}$	19.8	324
Tetracosane	$C_{24}H_{50}$	20.3	338
Pentacosane	$C_{25}H_{52}$	20.7	352
Esacosane	$C_{26}H_{54}$	21.3	366
Eptacosane	$C_{27}H_{56}$	21.8	380
Ottacosane	$C_{28}H_{58}$	22.4	394
Nonacosane	$C_{29}H_{60}$	23.0	408
Triacontane	$C_{30}H_{62}$	23.8	422
Hentriacontane	$C_{31}H_{64}$	24.6	436
Dotriacosane	$C_{32}H_{66}$	25.6	450
Tritriacosane	$C_{33}H_{68}$	26.8	464
Tetratriacosane	$C_{34}H_{70}$	28.1	478
Pentatriacosane	$C_{35}H_{72}$	28.8	492
Hexatriacontane (internal standard)	$C_{36}H_{74}$	31.7	506

2.5.2.3 Miliacin

The oven program for miliacin starts with a static of 1 minute at 100 °C, then a ramp at 30°C min^{-1} increases the temperature up to 315 °C, the program ends with a 12-minute static. Post run is at 315 °C for 15 minutes. The injection volume is 1 µL in splitless mode and the injector is set at 300 °C temperature. Helium flow is operated at 1.5 mL min^{-1}.

Under these conditions, the miliacin retention time is at 16.9 min and was integrated on the m/z ion 189, the control ion is the one with mass 425, that represents the molecular ion (miliacin molecular weight is 440 g mol^{-1}) subtracted of a methyl group. The internal standard's retention time was 9.1, the target ion for the integration was the one with m/z 245.

2.5.2.4 Sterols and stanols

Chromatography techniques such as liquid and gas chromatography for the determination of sterols and stanols are the most widely used. The abundance of isomers of sterols and stanols naturally present in environmental samples requires the use of highly sensitive detectors, hence mass spectrometry equipment guarantees an advantage when analyzing these proxies. Being slightly polar, sterols and stanols need to be derivatized: the polar hydroxyl groups must be transformed into nonpolar groups. The reaction occurs with the loss of the hydrogen of the hydroxyl group and its substitution with a non-polar group. In this project the derivatization was a silylation: the substitution takes place with trimethylsilyl group. The reaction guarantees easier volatilization of sterols and stanols necessary for further gas chromatography (GC) analysis.

For the analysis of sterols and stanols derivatization was necessary before the GC-MS analysis. Derivatization was performed using N,O-bis(trimethylsilyl)trifluoroacetamide with trimethylchlorosilane as catalyst (BSTFA + 1% TMCS, Merck, St. Louis, MO, USA). The reaction occurs with the loss of the hydrogen of the hydroxyl group and its substitution with a trimethylsilyl group.

100 µL of DCM and 100 µL of BSTFA + 1% TMCS were added to the dried samples, the liquid was transferred into a GC glass vial with a 250 µL glass insert and sealed with aluminum caps with pierceable septa. Vials were heated in a thermoblock for 1 hour at 70°C and analyzed after 24 hours. The heating treatment and one-day rest guarantee the best derivatization results (Battistel et al., 2015).

The analysis parameters are reported as followings. Temperature program was 150 °C, then 3 °C min^{-1} until 220 °C, then 0.7 °C min^{-1} until 275 °C, then 10 °C min^{-1} until 300°C held for 5 minutes. A post run of 15 minutes at 300 °C ended each run.

The m/z values for the SIM configuration are reported in Table 2.4.

Quantification of the analytes was obtained with cholesterol $^{13}C_3$ and corrected for the instrumental response factor.

A deep investigation of the fragmentation pathways of the fecal sterols and stanols is reported in Appendix A.

Table 2.4. Sterols and stanols retention times and SIM ions.

Analyte	Retention time (min)	Target ion (m/z)	Control ion (m/z)
Cholestanone	74.5	316	386
Coprostanol	70.1	215	370
Epi-coprostanol	72.3	215	370
Cholesterol	77.1	329	368
Cholestanol	77.8	215	445
24-Ethyl-coprostanol	83.5	215	398
Campesterol	83.8	343	382
24-Ethyl-epi-coprostanol	84.7	215	398
Stigmasterol	85.0	484	394
β-Sitosterol	87.5	396	357
Stigmastanol	88.2	215	473
Cholesterol $^{13}C_3$ (internal standard)	77.1	332	371

2.5.3 Analysis of the polar fraction

The preparation and analysis described subsequently is adapted from Barbaro et al., (2015) and Battistel et al., (2017).

The extraction of the polar fraction for the analysis of anhydrous sugars was performed with ASE 200 (Dionex Thermo Fisher Scientific, Waltham, MA, USA). ~1 g of soil sample was introduced into the stainless-steel cells where a glass fibre filter was placed at the bottom; internal standard was added (100 absolute nanograms of labelled Levoglucosan $^{13}C_6$ 98% CIL, Cambridge, UK), then cells were filled with diatomaceous earth (Thermo Fisher Scientific, Waltham, MA, USA). The extraction method consists of two static cycles at 100°C and 1500 psi of pressure using MeOH as solvent.

The extract was dried completely using Turbovap® (Caliper Life Science, Perkin-Elmer, Waltham, MA, USA) set at 35°C. The solid was redissolved in 500 µL of ultrapure water (ELGA LabWater, High Wycombe, UK) and filtered through 0.45 µm PTFE cartridges.

The analysis of sugars was performed through ionic chromatographer *Dionex*, ICS 5000 (Thermo Fisher Scientific, Waltham, MA, USA) coupled with a mass quadrupole detector MSQ Plus (Thermo Fisher Scientific, Waltham, MA, USA). The separation of the analytes is achieved using *CarboPac*™ MA1 column (athylvinylbenzene 55% e divinylbenzene, 2 x 250 mm, Thermo Fisher Scientific, Waltham, MA, USA), interferences due to ammino acids are removed with the use of a *AminoTrap*™ guard column (polymeric resin, 2 x 50 mm, Thermo Fisher Scientific, Waltham, MA, USA). Mobile phase solution is prepared by an eluent generator cartridge Dionex EGC III

NaOH (Thermo Fisher Scientific, Waltham, MA, USA). The flux is set at 250 µL min^{-1}. Injection volume is 50 µL.

The elution gradient of the sugars was the following: from 0 to 30 minutes the mobile phase concentration is 20mM, followed by a washing phase from 30 to 45 minutes (mobile phase concentration is 100 mM), the re-equilibration of the column is performed from 45 to 55 minutes using again a 20 mM mobile phase.

Mobile phase suppressor ASRS 500, 2 mm (Thermo Fisher Scientific, Waltham, MA, USA) was used to remove the mobile phase. Post-column the samples are added of a 0.025 ml min^{-1} flow of $MeOH/NH_4OH$ (7 ‰) solution; this is used to improve the analyte ionization by increasing the aqueous solvent. ESI electronspray is operated in negative mode and quadrupole is used in SIM mode. Table 2.5 reports the parameters of the sugars analyzed.

Table 2.5. Sugars retention times and SIM target ions.

Analyte	Retention time (min)	Target ion (m/z)
Levoglucosan	6.5	161
Mannosan	10.4	161
Galactosan	17.9	161
Levoglucosan $^{13}C_6$	6.5	167

2.6 Analysis of faunal dung

In order to build a dataset of fecal sterols and stanols from domestic animals, samples of fresh feces were collected. The animals considered the more frequent for human sustenance were donkeys, horses, sheep, goats, pigs, and cows. Excrements were collected in nature lodges or small farms, chosen to be more representative of animal life even in ancient times. Although the diet of the animals could have changed through the centuries (possibly affecting the relative concentrations of fecal sterols), the conditions of the sampling were chosen carefully. Particular attention was put on the life of the animals (outdoor and possibly in large fences) and their diet: farmers ensured the absence of antibiotics and animal feed was exclusively fresh grass, hay, and cereals. For each animal species, four excrement samples were collected from at least four different farms to ensure variability within the same animal group. To evaluate the

possible degradation within the first days, “ready-made” and dry samples were collected. Table 2.6 reports the specifications for each sample collected. Samples were collected into plastic tubes with caps and stored in freezer (-20 °C) until freeze-drying.

Table 2.6. Animal dung collection information.

Farm	Animals	Food	Number of samples
Farm House (VE)	Sheep	Grass, hay	5
	Horses	Grass, hay, cereals (corn/puffed oat)	4
Lo Stralisco nature lodge (TV)	Goats (4), donkeys (4)	Hay	8
L’impronta nature lodge (VE)	Sheep (4), goats (4), donkeys (4)	Grass, hay	12
Il Rosmarino nature lodge (VE)	Sheep (4), horses (4), goats (4), donkeys (4)	Grass, hay	16
La chioccia agritourism (VE)	Sheep (4), goats (4), donkeys (4)	Hay, alfalfa, corn	12
Villa Condulmer stables (TV)	Horses	Hay, cereals	4
Ca’ Donadel agritourism and nature lodge (TV)	Cows	Hay, grass, proteins of soy, sorghum.	4
	Pigs	Corn, soy, barley, whey	4
Case Bianche stables (VE)	Horses	Hay, oat	5
	Sheep (2), donkeys (2)	Grass, hay	4
Al vecio moraro agritourism (TV)	Pigs	Bran, barley, soy, vegetables	4
Bi-Farm nature lodge (VE)	Cows	Grass, hay	4
Scattolin farm (VE)	Cows	Hay, corn	4
	Pigs	Barley, alfalfa, corn, hay	4
De Faveri agritourism (VE)	Cows	Alfalfa, hay, corn	4

Samples were freeze-dried for a maximum of 48 hours after collection. Freeze samples were opened and put into the high vacuum freeze drier Lio5P (5Pascal, Trezzano sul Naviglio, MI, IT) at a temperature of -50 °C and a pressure of 6 mbar; cycles of 24 hours were performed. After each cycle the samples were moved to the next step or underwent another cycle.

Freeze-dried samples were milled into an automatic mortar.

The extraction of fecal sterols and stanols was performed putting 0.5 g of sample into a glass flask with cap (volume 50 mL) and adding 2.0 µg absolute of internal standard. Extraction was performed with two cycles using 20 mL of DCM each time and 10 minutes in the ultrasonic bath. The ~40 mL of liquid extract was collected into glass vial, and 2 g of anhydrous sodium sulphate was added to the liquid to avoid humidity traces and stored in the fridge (4 °C).

A 5 mL aliquot was purified using Pasteur pipette filled with glass-wool and activated magnesium silicate (Florisil®, 100-200 mesh, Merck, St. Louis, MO, USA). 1 mL of

the purified extract was transferred into a small volume glass vial and dried under a gentle flux of nitrogen and stored in the fridge until analysis.
Derivatization and analysis were analogue to those of the DE soil samples, for details see sections 2.5.1 and 2.5.2.4.

2.7 Statistical analysis and modelling

The pool of data was analyzed using R-studio (version 4.2.1) base packages (version 3.5.1) for statistical hypotheses testing. Pearson correlation coefficients, Wilcoxon-Mann-Whitney and Student's t-tests were used to evaluate the relationship between all the variables treated. The Shapiro and Wilk test was used to check the normality of distributions. Variables were compared using the t-test or, when the normality hypothesis was not satisfied, using the non-parametric Wilcoxon-Mann-Whitney test.
The packages necessary for building the models and to test their performances were the following: for linear discriminant analysis (LDA), package `MASS` was employed, for classification tree setup and figures the packages `tree` and `rpart.plot`, respectively, were used, to perform random forest (RF) models the package `randomForest` was used.

Results and Discussion

The results and the related discussion of the analysis performed are presented in the following chapters. Chapter three is focused on soil morphology with the analysis of radiocarbon dating, metallic elements, and carbon. Chapter four is dedicated to the biomarkers related to the reconstruction of biomass burning and environmental conditions. Lastly, Chapter five includes the fecal proxies results. This latter is divided into two parts: the first concerns the results obtained by the analysis of the animal feces and the attempts to find a classification model for the sources identification, whereas the second part is focused on the fecal sterols and stanols content in DE samples.

The results and discussion presented in chapter three were part of a previously published in: Bortolini, M.*, Agnoletto, F. C., Argiriadis, E., Nicosia, C., McWethy, D. B., Devos, Y., Stortini, A. M., Baldan, M., Roman, M., Vendrame, T, Scaggiante, R, Bruno, B, Pojana, G., Battistel, D. Insight into the carbonaceous fraction of three cultural layers of different age from the area of Verona (NE Italy), CATENA, 217 (2022), (DOI: 10.1016/j.catena.2022.106453).

3 Soil morphology and microscopic features

The morphology and the microscopic characteristics of soil are key features for understanding the dynamics underlying the formation of DEs.

As reported previously (section 1.1), the presence of burned particles and anthropogenic remains has been investigated as a direct consequence of human presence in a site and is one of the main features used to reconstruct the dynamics at the base of the deposition of cultural layers. Moreover, the presence of charred particles contributes to the characteristic color of the DE layers. In this chapter, a detailed characterization of the carbonaceous fraction is reported, discussed, and correlated with other proxies including soil carbon content and trace elements.

3.1 Radiocarbon dating

The VP I, VP II, SPM I and SPM II profiles result very close together in terms of location (Figure 2.1) and as such they belong to a correlated DE deposit. Following the indication proposed in Pessenda et al., (2021), the humin fraction of the soil was initially dated as it is, differently from charcoal fragments, potentially less affected by post-depositional processes. Conventional Radiocarbon Ages (CRA) and calibrated ages are reported in Table 3.1 and Figure 3.1.

The radiocarbon calibrated ages obtained from the humin fraction span a period between 75 and 380 AD (significance corresponding to 2σ). These radiocarbon ages did not provide a chronological sequence, suggesting that the matrix has been mixed and reworked. Indeed, radiocarbon ages do not increase with depth, and this is independent of the typology of the fraction of carbon used (i.e., charcoal or humin). A similar sequence of mixed-aged materials was observed in DE from Brussels (Devos et al., 2017) and Florence (Nicosia, 2018).

Table 3.1. Conventional Radiocarbon Ages and calibrates ages obtained from the humin fraction and charcoal fragments in SPM II.

Profile	Aliquot	Material	Lab. code	CRA (yr BP)	Calibrated date AD
SPM II	1 top	humin	580606	1890 ± 30	76-232 (95.4%)
	2 top	humin	580607	1830 ± 30	126-253 (84.4%)
					290-320 (11.0%)
	3 top	humin	580608	1760 ± 30	234-380 (95.4%)
	4 top	humin	580609	1890 ± 30	76-232 (95.4%)
	5 top	humin	580610	1840 ± 30	124-250 (91.6%)
					294-310 (3.8%)
	1 top	charcoal	586538	1820 ± 30	152-256 (69.3%)
					284-326 (22.2%)
	3 top	charcoal	586539	1600 ± 30	416-545 (65.4%)
	5 top	charcoal	586540	1740 ± 30	245-402 (95.4%)

Some of the hand-picked charred fragments from the SPM II profile were radiocarbon dated to test the hypothesis of post-depositional events. Calibrated ages obtained from charcoal ranged from 150 to 545 AD. Small differences between the radiocarbon dating of charred material versus humin fraction were found. These differences could be due to the diverse processes of formation of the analyzed substances. Humic carbon represents a part of the matrix originally deposited (van der Plicht et al., 2019; van Mourik et al., 2012) and then enriched in coarser fragments and charcoal during the subsequent centuries. Indeed, the humin fraction is also defined as allochthonous organic carbon, that is the carbon fraction deposited in the matrix prior to the formation of the autochthonous carbon, more easily oxidable, in soil (van der Plicht et al., 2019). The period of formation of DE in VP and SPM is nearly in line with the one exposed in the nearby urban excavation of Via Mazzini 41, dating to the late 4th – early 6th century AD (Bruno and Fresco, 2014). The dates obtained at SPM show that in DE studies it is crucial to define the context of the material being dated. Dated material from an in-situ DE or cultural layer dates – roughly– its moment of formation (with all limitations from old wood effect, stocked wood, re-use, heirlooms, etc.). Dated material from dumped or remobilized DE or cultural layers provides us with a *terminus post quem* with respect to the moment in which those sediments were dumped. Hence, from our dating results it is possible to consider radiocarbon dating not only a base technique to attribute the soil to an exact time period, but also a starting point to draw the succession of different anthropic events.

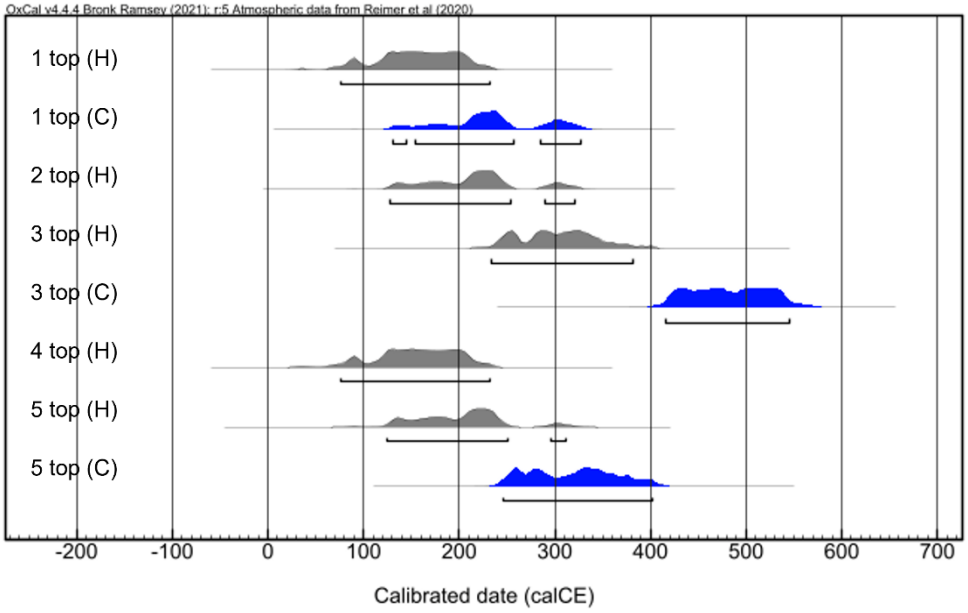

Figure 3.1. OxCal plot of the radiocarbon dating results. Uncertainties are expressed with a 95.4% of significance. Ages refers to common era (CE) calendar calibration. Samples are referred to the SPM II profile, (H) and grey curves represent the humin fraction, (C) and blue curves are referred to the charcoal fraction dating.

3.2 Soil Micromorphology

In all the sites, a specific focus is given to the early formation of the DE deposits with respect to the underlying substrate. Second, the degree of expression of syn- and post-depositional processes and transformations that are normally considered as key in DE formation (i.e., bioturbation, decalcification, cultivation, etc.) is established. For the detailed description of the thin sections, please refer to Appendix B.

For the site of Via Pigna (VP), the base of the sampled sequences is given by a mortar floor (Figure 3.2a), a typical feature of early medieval dwellings that re-occupied the remains of former Roman structures (Nicosia, 2018). In profile VP I (sample 3), above this floor, there is a 1-cm thick remnant of 'trample' (*sensu* Banerjea et al. 2015). This micromorphological characteristic is consistent with sediments that gradually accumulated in situ on this type of floor. These deposits are composed of compacted loam-textured material, particularly rich in wood ash, feces (carnivore/omnivore and herbivore), bone fragments, charcoal, and horizontally-lying articulated phytoliths.

These sediments are markedly different from the DE lying above, as suggested by the sudden boundary between the two (Figure 3.2a). This indicates a rapid change in the sedimentation process, from a trampled domestic context to a fast-growing DE, derived from dumping or backfilling. The DE is composed of poorly sorted clay loam with randomly distributed archaeological components (charcoal, abundant metal slags and droplets, ceramic, and fecal material). The slags exhibit a variety of compositions: fayalite (Fe_2SiO_4) in larger or smaller crystals, portions of amorphous glass with vesicular porosity, amorphous Fe-oxides or dendritic wüstite (FeO). Bioturbation in this DE layer is weakly attested and there are no traces of human reworking (i.e., no traces of slaking derived from agricultural practices).

The DE found in vicolo San Pietro in Monastero (SPM) lies directly above a layer of levelled construction rubble with a clay-loam matrix, containing Roman-age ceramic. The DE collected at SPM contains a randomly distributed admixture of waste from domestic activities (ash, ceramic and bone fragments, carnivore-omnivore excrements, eggshell) with lower quantities of metal slags and droplets from metalworking compared to VP. The matrix is the typology of fine-earth matrix (i.e., < 2 mm fraction). Bioturbation is present but not very expressed, some biogenic channels and earthworm granules are indeed present. Yet, the total biogenic fabric of calcareous mull-like DE horizons, observed even in the nearby site of Verona-Via Mazzini 41 (Nicosia 2018) and considered to be one of the key forming factors of DE (Macphail, 1994, 1981; Sukopp, 1979) are here absent. In SPM profiles there is no evidence of hiatuses in sedimentation, cultivation, or to post-depositional processes such as decalcification.

Soil bioturbation is the main process observable in the Bronze Age cultural layer of Castelar di Leppia (CDL). This is evidenced by the expressed channel microstructure, by the presence of crescent-shaped complete infillings ('bow-like features' *sensu* Bullock et al., 1985), loose discontinuous infillings of mineral fecal pellets, larger coalescing earthworm droppings, and by scattered earthworm granules. The fine matrix of the cultural layer results enriched in silt-sized calcareous material, giving rise to a crystallitic b-fabric, and in black particles. This contrasts with the local substrate, which is a partially decalcified Cambisol formed on the Adige river sediments (De Zanche et al., 1977). The enrichment in calcareous fine material is probably the outcome of mixing processes with sediments derived from lower horizons of this Cambisol, which are still very calcareous. Such mixing derives from the fact that deeper calcareous

sediments were dug up (i.e., pits, wells, quarrying of earth material, etc.). The occurrence of black particles probably derives instead from the admixture to the cultural layer of combustion by-products (ash, charcoal, soot, etc.) which have been strongly reworked by bioturbation.

Figure 3.2 shows: (1) Intimately commixed black particles and mineral material in DE composed of waste from proximal metallurgical activities (VP I; sample 3, Figure 3.2b, c) (2) finely (< 25 um) comminuted and distanced black particles, some of which still retain the structure of vegetal tissues in DE derived from ash-rich domestic waste disposal (SPM, sample 12, Figure 3.2d, e); (3) very finely comminuted black particles in a cultural layer that underwent longer bioturbation (CDL, sample 1, Figure 3.2f, g).

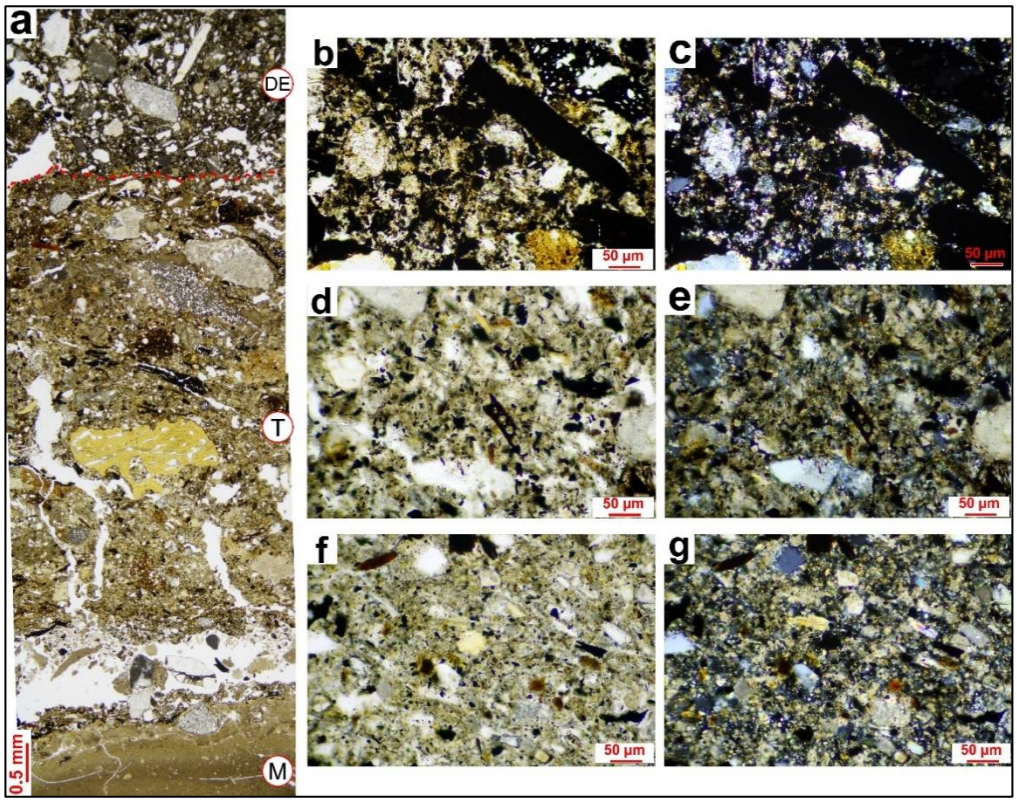

Figure 3.2. (a) VPI, sample 1. Contact between mortar floor (M), trample layer (T) and Dark Earth (DE). Note the compositional difference and the abrupt limit (dotted line) between trample accumulated *in situ* and DE resulting from dumping of waste. Note the horizontal alignment of coarser fabric units (bone, rock fragments, charcoal) in layer T. PPL. (b, c) VP I; sample 3. Black particles viewed at 100x with substage condenser inserted. PPL and XPL. (d, e) SPM I, sample 6. Same as above. (f, g) CDL, sample 1. Same as above.

3.3 Metallic droplets and trace elements

During the microscopy observation of the samples for charcoal analysis (see section 3.5), many spherical inclusions were detected in both VP and SPM. A deeper investigation of the distribution of these droplets showed that they were most abundant in the > 63 μm fraction as reported in Table 3.2.

Table 3.2. Distribution of the ferrous microspheres of different size from the three sites Via Pigna (VP) vicolo San Pietro in Monastero (SPM) and Castelar di Leppia (CDL).

Site	500 μm >d>250 μm	250 μm >d>125 μm	>63μm
VP	2 ± 1 counts g^{-1}	15 ± 9 counts g^{-1}	30 ± 10 counts g^{-1}
SPM	-	-	7 ± 1 counts g^{-1}
CDL	-	-	-

A representative selection of metallic droplets was handpicked and analyzed by SEM-EDS to study their morphology and elemental composition. A typical SEM micrograph is shown in Figure 3.4. Such droplets exhibit a regular spherical shape with a diameter ranging between 80 and 280 μm. The elemental analysis evidenced that they are mainly composed of Fe, Al, and oxygen (coherently with the phases identified during the thin section observation). Ferrous spheres in modern soils have been usually associated to fly ash from metallurgical activities and road dust pollution (Bourliva et al., 2016; Fisher et al., 1978; Howard and Orlicki, 2016; Ngu et al., 2007). To the best of our knowledge, the presence of this kind of inclusion has never been reported before in the literature for DE. The lower quantity of metallic droplets, all belonging to the smallest dimensional class, in SPM than in VP, suggests the latter as closer to the source site of the metallic particles, transported to the neighboring site SPM.

The presence of metallic droplets prompted to investigate in detail the elemental composition of the soils for metals and other elements typically associated to anthropogenic soils, such as P. Two samples from each profile (VP I, VP II, SPM I and SPM II) were analyzed together with one sample from CDL. The concentrations are reported in Table 3.3, expressed as enrichment factors (EF) as proposed by Reimann and De Caritat, (2000).

Table 3.3. Mean values of the Enrichment Factors from two sections at vicolo San Pietro in Monastero (SPM I and SPM II) via Pigna (VP I and VP II), and from Castelar di Leppia (CDL).

Site	n	EF Fe	EF P	EF Ca	EF Mg
SPM I	2	2.6 ± 0.1	3.6 ± 0.3	30 ± 1	4.1 ± 0.2
SPM II	2	2.6 ± 0.1	3.8 ± 0.1	25 ± 3	4.2 ± 0.4
SPM	4	2.6 ± 0.1	3.7 ± 0.1	28 ± 2	4.2 ± 0.2
VP I	2	10 ± 5	7 ± 1	23 ± 1	5 ± 1
VP II	2	20 ± 13	9 ± 3	40 ± 17	5.9 ± 0.2
VP	4	14 ± 6	8 ± 1	30 ± 8	5.0 ± 0.5
CDL	1	2	3	18	5

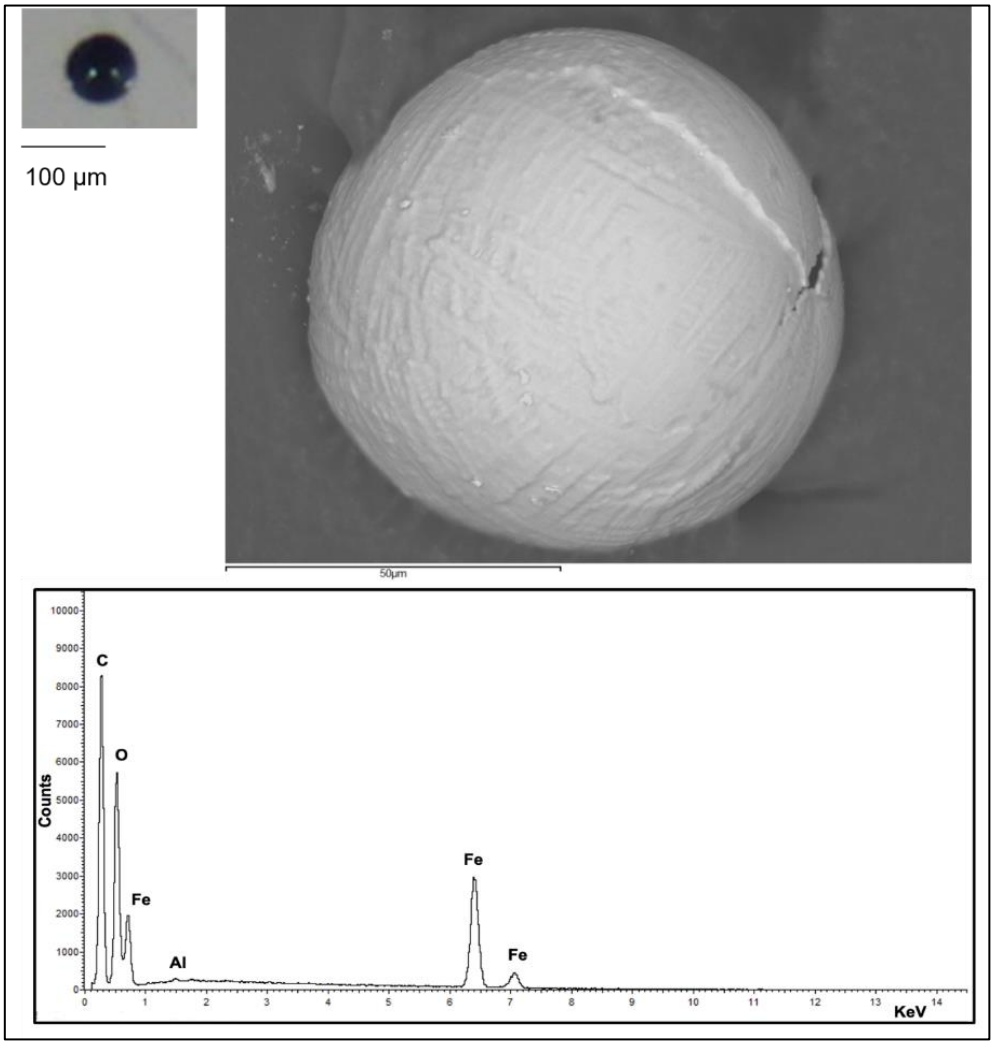

Figure 3.4. Stereomicroscopic image of ferrous sphere isolated from VP site; SEM image collected at 1.5k magnification. Punctual EDS spectrum and element parameters and concentrations detected.

Sutherland (2000) proposed that an elemental concentration should be considered anomalous when its EF exceeds a value of ~20. According to this threshold level, all examined DE profiles were highly enriched in P and Ca, while VP was also significantly enriched in Fe. The high value of iron EF in VP is in accordance with the higher presence of metallic droplets found in VP with respect to the other sites. Although the EF evidenced marked differentiation between the three sites, they show similar values within the profiles of the same site. This feature further supports the occurrence of a horizontal rather than vertical variability in DE deposits.

The strong presence of P could possibly entail a phenomenon of downward seeping, identified by Macphail and Linderholm, (2004). Nevertheless, the concentration of P found in DE is in agreement with the results obtained in similar ancient anthropogenic soils in northern Italy (Migliavacca et al., 2013; Vittori Antisari et al., 2013). In fact, the high concentrations of P (and Ca) determined in other past settlement soils are attributed to the presence of human waste (i.e., plant and animal material, ash) being incorporated into the soil matrix (Entwistle et al., 2007, 2000; Pierce et al., 1998). In undisturbed soil patches, the concentration of P is reported to be much lower than in the vicinity of settlements; additionally, the increase in the amount of this element depends on the continuity of human occupation in a site. Moreover, it has been reported that the persistence of P makes it still detectable hundreds of years after the fertilization has ended (Dambrine et al., 2007; Entwistle et al., 1998; Guttmann et al., 2008; King, 2008; Leonardi et al., 1999; Linderholm, 2003; Vittori Antisari et al., 2013; Wells et al., 2000).

Calcium EF shows a particular increase in VP and SPM, with respect to CDL. Ca and Mg can derive from different sources in anthropogenic soils: in the examined sites calcium and magnesium can originate primarily from calcareous-calcarenites minerals, present in the soil matrix. These could then be transformed into the soil thanks to the continuous supply of organic material which, in turn, undergoes humification processes. Such processes affect pH of the soil, determining favorable conditions for the dissociation of carbonates and the segregation of the ions by organic matter (Zamanian et al., 2016, Rowley et al., 2018; Shi et al., 2017). Ca and Mg can also come from phosphatic minerals, also suggested by the significative presence of P. Therefore, Atkinson et al. (2010) Lehmann and Stahr (2007) and Steiner et al. (2007) suggested the increasing of concentration of micronutrients in soil as due to a strong availability of biochar and charcoal.

3.4 Carbon

The total organic carbon (TOC) and total inorganic carbon (TIC) were determined together with the recalcitrant organic carbon (ROC). The concentration profiles along all sites did not exhibit a particular trend, concentrations of each sample are reported in the table in Appendix C. Therefore, the mean concentrations of TOC, ROC and TIC for each profile were considered, as reported in Table 3.4. The highest ROC and TOC mean values were detected in VP for both the profiles VP I and VP II, while SPM showed a higher TIC.

Table 3.4. Mean values organic (TOC), recalcitrant (ROC) and inorganic carbon (TIC) from two sections at vicolo San Pietro in Monastero (SPM I and SPM II) via Pigna (VP I and VP II), and from Castelar di Leppia (CDL).

Profile/site	n	TOC (%)	ROC (%)	TIC (%)
SPM I	9	1.3 ± 0.1	0.8 ± 0.1	4.3 ± 0.2
SPM II	10	1.3 ± 0.2	0.7 ± 0.1	4.2 ± 0.2
SPM	19	1.3 ± 0.1	0.7 ± 0.1	4.3 ± 0.1
VP I	6	3.3 ± 3.0	1.6 ± 0.7	3.2 ± 1.2
VP II	6	2.5 ± 1.7	1.5 ± 0.7	3.0 ± 1.5
VP	12	2.9 ± 1.4	1.5 ± 0.4	3.1 ± 0.8
CDL	3	0.9 ± 0.4	0.5 ± 0.1	3.6 ± 0.1

The means were compared using the t-test or, when the normality hypothesis was not satisfied, using the non-parametric Wilcoxon-Mann-Whitney test. The resulting p-values of the tests are reported in Table 3.5. As shown, TOC, ROC and TIC are not significantly different between parallel profiles from the same sites. Nevertheless, VP has significantly higher values of TOC and ROC when compared to both SPM and CDL, as well as a depletion in TIC. However, CDL shows significantly lower TOC and ROC than SPM.

Table 3.5 Comparison between mean values of organic (TOC), recalcitrant (ROC) and inorganic carbon (TIC). p-values obtained using the Student's t or the Wilcoxon-Mann-Whitney. p-values< 0.05 (underlined) indicate a significant difference between the means.

	TOC	ROC	TIC
SPM I vs SPM II	0.91	0.27	0.65
VP I vs VP II	0.55	0.78	0.74
SPM vs VP	0.033	0.001	0.008
VP vs CDL	0.009	$9 \cdot 10^{-5}$	0.18
SPM vs CDL	0.008	0.0003	$4 \cdot 10^{-8}$

Mean values of TOC range between 0.9 and 3.3% and are consistent with the concentrations found in other DE horizons all over Europe. Devos et al. (2013, 2017), reported a TOC content between 0.5 and 3.7% in Brussels, while slightly lower organic carbon concentrations were observed in Padua (0.7-1.9%) and in Florence (0.25-1.25%) (Nicosia et al., 2012, 2019). These values were ascribed to the result of handling processes (domestic waste or byre) and manure addition to the soil (Devos et al., 2017; Yannick Devos et al., 2013a; Wiedner et al., 2015).

Although the three sites are significantly different when compared for TOC, most variability is due to the ROC concentrations, ranging between 0.5 and 1.6% (Table 3.4). A comparison of ROC between the sites returns the lowest p-values, following the same relationship observed for TOC as ROC concentration resulted VP > SPM > CDL (Table 3.4).

It must be noted that TOC and ROC values are significantly higher in VP than in SPM despite the vicinity of the two sites. This suggests that the horizontal variability in DE, likely related to different practices occurred is significant even for short-range distances.

3.5 Charcoal

The charcoal content was determined following the procedure proposed by Whitlock and Larsen, (2002). The concentration of charcoal pieces along all the sites did not show any trend; the amounts of charcoal counted in each sample are reported in Appendix C.

The results of the charcoal frequency analysis are reported in Table 3.6 as mean values and corresponding confidence intervals (95%) for each profile and site, separated into five dimensional classes (1 mm, 500, 250, 125 and 63µm). As shown, the number of charcoal fragments increases with decreasing the dimensional class, for all profiles and sites. The dimensions of the charred particles are widely used as indication of the proximity of a fire event (McWethy et al., 2010; Whitlock and Larsen, 2002).

Table 3.6. Mean values of the charcoal counts g^{-1} in samples sieved at 1000, 500, 250, 125 and 63 µm from two sections at vicolo San Pietro in Monastero (SPM I and SPM II) and via Pigna (VP I and VP II), for the whole sites (SPM and VP), and from Castelar di Leppia (CDL).

Site	N	CHAR1000	CHAR500	CHAR250	CHAR125	CHAR63
SPM I	27	NA	NA	2 ± 1	11 ± 4	17 ± 4
SPM II	30	NA	NA	3 ± 2	9 ± 3	15 ± 3
SPM	57	NA	NA	3 ± 1	10 ± 2	16 ± 2
VP I	18	2 ± 1	3 ± 1	20 ± 8	50 ± 20	60 ± 30
VP II	18	0.5 ± 0.5	2 ± 1	16 ± 7	25 ± 6	45 ± 20
VP	36	1.1 ± 0.6	2.4 ± 0.9	18 ± 5	40 ± 10	50 ± 20
CDL	9	NA	NA	0.4 ± 0.4	13 ± 6	80 ± 45

Non-parametric Wilcoxon-Mann-Whitney test was conducted to compare charcoal fingerprint distribution between sites and profiles. The p-values reported in Table 3.7 support the hypothesis that, similarly for TOC, ROC and TIC concentrations, significant inter-site rather than intra-site differences exist. VP shows the highest abundance of charcoal fragments in all dimensional classes, except for the smallest (63 µm). In addition, VP is the only site where coarser charcoal particles (> 500 µm) were detected. CDL exceeds only in CHAR63 with respect to SPM, but this dimensional class is not significantly different from the corresponding one in VP. The same dimensional differences were highlighted in section 3.2, suggesting alternative processes involved in DE formation. Further information about the charcoal inner structure were obtained by µ-Raman spectroscopy performed on the surface of the largest charcoal fragments.

Table 3.7. P-values for the comparison between mean values of charcoal counts in samples sieved at 1000, 500, 250, 125 and 63 µm. p-values obtained using the Student's t or the Wilcoxon-Mann-Whitney test. p-values < 0.05 (underlined) indicates a significant difference between the means.

	CHAR1000	CHAR500	CHAR250	CHAR125	CHAR63
SPM I vs SPM II	-	-	0.97	0.41	0.38
VP I vs VP II	0.053	0.89	0.71	0.22	0.80
SPM vs VP	$\underline{10^{-7}}$	$\underline{10^{-11}}$	$\underline{10^{-11}}$	$\underline{10^{-9}}$	$\underline{10^{-8}}$
VP vs CDL	$\underline{0.039}$	$\underline{10^{-3}}$	$\underline{10^{-5}}$	$\underline{10^{-3}}$	0.059
SPM vs CDL	-	-	$\underline{0.050}$	0.25	$\underline{10^{-6}}$

3.5.1 Raman spectroscopy

Tuinstra and Koening (1970) reported a comparison between the Raman spectra of graphite (crystalline) and graphite-like (nanocrystalline/amorphous) materials. The spectrum of crystalline graphite is defined by only one band at 1575 cm^{-1}; moving to

the less crystalline carbon material, a structure at 1355 cm^{-1} appears. The intensity of this latter increases with the increase of unorganized carbon domains and with the decrease of the graphite crystal size. Comparing the spectra of these materials with that of diamond (where only one line at 1332 cm^{-1} is present) the authors hypnotized the presence of a certain quantity of tetrahedrally bonded carbon atoms in the disordered charcoal particles (Tuinstra and Koenig, 1970). Therefore, the spectrum of non-crystalline carbon materials is characterized by two features: the G (graphite-like) band around 1580-1600 cm^{-1} and the D (diamond-like) band cantered around 1350 cm^{-1}. These two are related to different Raman active modes in graphite (G band) and amorphous carbon (D band) (Schwan et al., 1996; Tuinstra and Koenig, 1970).

Charred materials are associated with the presence of ordered regions of graphite crystals, thus the D and G bands have been related to two different vibrational modes of benzene-correlated materials. Although the scientific community agrees on the assignment of the G band to the vibration of the symmetrical E_{2g} mode (Ferrari and Robertson, 2000; Jawhari et al., 1995; McDonald-Wharry et al., 2013; Tuinstra and Koenig, 1970; Wang et al., 1990) (Figure 3.5), the phonon correlated to the D band is still debated and less easily interpretable. Some researchers assigned the D band to the vibration of the A_{1g} mode (Ferrari and Robertson, 2000; Jawhari et al., 1995; Sheka et al., 2020; Trusovas et al., 2016; Tuinstra and Koenig, 1970; Zhang et al., 2015, among the others) (Figure 3.5), while other scholars are inclined to consider this feature as the result of a mechanism of double resonant Raman scattering (Lambrecht et al., 2021; Saito et al., 2002; Thomsen and Reich, 2000; Zickler et al., 2006).

The degree of 'graphite-like' and 'diamond-like' features of carbon materials is related to the amount of sp^2 and sp^3 hybridization state of the carbon atoms (Ferrari and Robertson, 2004, 2000; Ishimaru et al., 2007b; Nemanich and Solin, 1978). Amorphous carbon can contain various ratios of sp^2 and sp^3 orbitals, arising from different conditions involved in the burning and aging processes, as well as from the different biomass precursors that can be burned. The sp^3 hybridization is related to σ bonds, whereas the possible state for sp^2 orbitals is the π state. In the σ bonds, the electron disposition geometry is two-centre orbitals between adjacent atoms, and it is a structure related to a short-range order; whereas conjugated electronic structures are specific to π bonds, in which the hybridization cannot be defined by only one atomic bond. In the π hybridization, each bond contains the contribution of other bonds and the structure is defined by a medium-range order (Ferrari and Robertson, 2000). Differences in the concentration of carbon atoms having a sp^2 hybridization arise from structural changes

involving a modification from ordered domains, typical of graphitic materials, to more amorphous ones, such as glassy carbon (Ferrari and Robertson, 2004). Ferrari and Robertson (2004) report also a continuous increase of the G band full-width half maximum (FWHM) as the disorder increases. This relation is also supported by Ishimaru et al., (2007), who associated the decreasing of the FWHM of the G band with the development of carbon crystallites. Ishimaru et al., (2007) linked the G and D band to *"the in-plane vibrations of sp^2-bonded carbon with structural imperfections (D band for disorder) and the in-plane vibrations of sp^2-bonded crystalline carbon (G band for graphite)"*. They also identified a band between 1190 and 1120 cm^{-1} due to the presence of bonded sp^3 carbon atoms. These sp^3 carbons are interpreted as the cross-linking between adjacent crystallites, interrupting the ordered structure of carbon crystallites. The formation of this structure is related to the heating treatment responsible for the charring process (Ishimaru et al., 2007b).

In general, any blueshift (increasing of the Raman shift position) of the G band, as well as the decrease in the ratio between the intensities of the valley between D and G band (I_V, Figure 3.6) and the G band intensity (McDonald-Wharry et al., 2013; Ribeiro-Soares et al., 2013) (I_V/ I_G) are correlated with the increase of the charring temperature (Figure 3.5), Specifically, the I_V/ I_G decrease is related to the removal of the molecular oxygenated defects (Inoue et al., 2017). With the increase of the charring temperature, a decrease in the D band position is also observable (redshift), and it is due to the removal of the sp^3 amorphous structures (McDonald-Wharry et al., 2013; Smith et al., 2016).

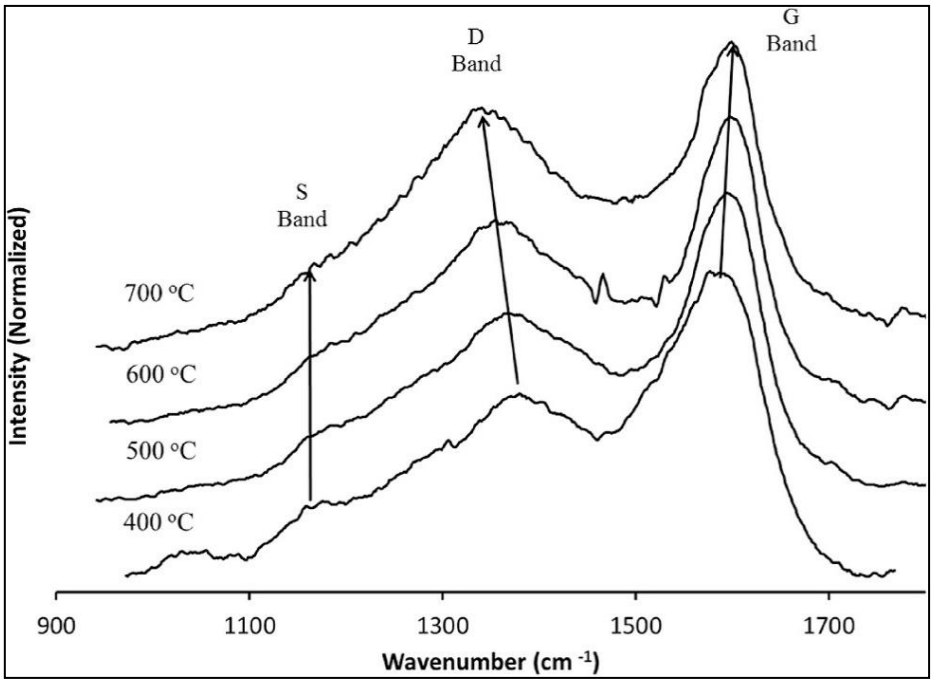

Figure 3.5. Raman spectra in which the shifts of the bands are correlated with the increasing temperature of carbonization (Smith et al., 2016).

Figure 3.6 reports the typical Raman spectra collected on three fragments of wooden charcoal coming from the sites studied in this work and showing the characteristic D and G bands (~1350 and ~1600 cm^{-1}, respectively), as obtained after deconvolution (Cohen-Ofri et al., 2006; De Sousa et al., 2020; Dennison and Holtz, 1996; Dresselhaus et al., 2005; Ishimaru et al., 2007a; Jawhari et al., 1995; Kawakami et al., 2005; Tuinstra and Koenig, 1970).

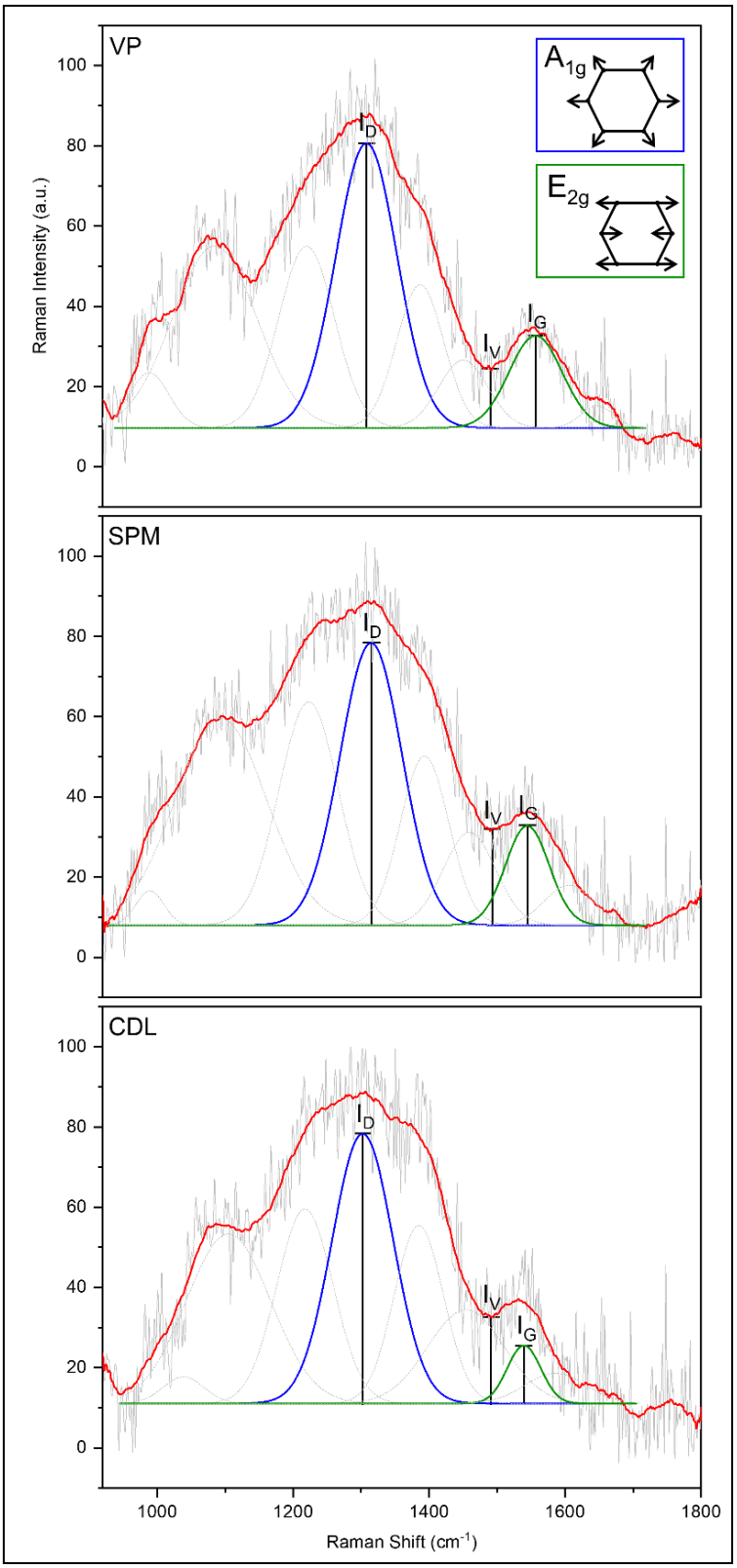

Figure 3.6. Example of three typical Raman spectra acquired on samples from the of interest for in this study. Fit (red line) to the VP, SPM and CDL spectra through deconvolution (thin lines) into the identified spectral components. The bands attributed to the D and G components and used for elaborations and interpretation are highlighted in blue and green, respectively. Top window: sketch showing the attribution of the phonons of the D and G bands.

The main diagnostic parameters obtained from Raman spectra for all the profiles and sites are summarized in Table 3.8, while the values for each sample are reported in the Appendix C. Table 3.8 reports the mean values and the confidence interval (95%) of the positions (cm^{-1}) of D and G bands, the ratio between the intensities of the two bands (I_D/I_G), I_V/I_G, and the in-plane crystallite size (L_a) of charred particles in nm. This latter was calculated using the equation proposed by Cançado et al. (2006), in which I_D/I_G is inversely correlated to the dimension of the graphite crystallites (L_a). This relation was firstly observed by Tuinstra and Koenig (1970) and then described with an empirical formula to calculate L_a from the I_D/I_G measured via Raman spectroscopy employing a laser wavelength of 514.5 nm (Knight and White, 1989; Mernagh et al., 1984). The equation was initially applicable only to the spectra acquired with the same laser, while Cançado et al., (2006) proposed a new empirical and generalized formula (4):

$$L_a(nm) = (2.4 \times 10^{-10})\lambda_l^4 \left(\frac{I_D}{I_G}\right)^{-1} \quad (4)$$

where λ_l is the laser wavelength, to derive L_a from Raman spectra acquired under any laser condition. The equation describes the strong dependency on the laser energy not only for the position of the D band, but also for the ratio between the two bands intensities.

Table 3.8. Mean values ± confidence interval (95%) of the parameters from Raman spectra obtained in charcoal samples from two sections at vicolo San Pietro in Monastero (SPM I and SPM II) via Pigna (VP I and VP II), and from Castelar di Leppia (CDL).

Site	n	Band G (cm^{-1})	Band D (cm^{-1})	I_D/I_G	I_V/I_G	L_a (nm)
SPM I	4	1557 ± 6	1314 ± 9	1.7 ± 0.3	0.89 ± 0.07	52 ± 8
SPM II	12	1554 ± 2	1314 ± 4	1.7 ± 0.1	0.94 ± 0.04	54 ± 3
SPM	16	1555 ± 2	1314 ± 3	1.7 ± 0.1	0.93 ± 0.03	53 ± 3
VP I	6	1549 ± 6	1318 ± 3	1.9 ± 0.5	0.94 ± 0.11	52 ± 10
VP II	4	1553 ± 4	1316 ± 18	1.9 ± 0.4	0.89 ± 0.28	47 ± 10
VP	10	1551 ± 4	1317 ± 5	1.9 ± 0.3	0.91 ± 0.10	50 ± 8
CDL	3	1542 ± 5	1316 ± 9	1.7 ± 0.4	1.16 ± 0.18	53 ± 10

The L_a values obtained in this work have an average dimension of about 50 nm, slightly higher than the Amazonian analogous (Jorio et al., 2012; Ribeiro-Soares et al., 2013) and more similar to the graphitic materials studied by Cuesta et al. (1998). A statistical comparison between the parameters reported in Table 3.8 was carried out using the Student's t-test (or Wilcoxon-Mann-Whitney test). The p-values of the test performed within and between the sites are reported in Table 3.9.

Table 3.9. Comparison between mean values of Raman Parameters. p-values obtained using the t-Student or the Mann-Whitney test depending on the normality of the distribution previously checked with the Shapiro-Wilk test. p-values < 0.05 (undelined) indicates a significant difference between the means.

	Band G (cm^{-1})	**Band D (cm^{-1})**	**I_D / I_G**	**I_V / I_G**	**L_a (nm)**
SPM I vs SPM II	0.13	0.68	0.66	0.10	0.67
VP I vs VP II	0.31	0.82	0.78	0.58	0.55
SPM vs VP	0.08	0.077	0.19	0.76	0.34
VP vs CDL	0.012	0.46	0.24	0.005	0.42
SPM vs CDL	0.006	0.48	0.96	0.019	0.98

The structure of charcoal fragments coming from the two urban sites (VP and SPM) and their profiles does not show significant differences. Conversely, the older site CDL exhibited markedly higher I_V/I_G and lower G-band position values.

3.6 Correlations among microscopic variables

For a deeper investigation of the meaning of the variables investigated so far in the context of DE, Pearson correlation coefficients between pairs of variables are reported in Table 3.10. As shown, TOC and ROC were significantly and positively correlated (r = 0.60, p-value <0.01). Qi et al. (2017) found similar positive correlations between TOC and ROC in agricultural and parkland soils, suggesting a possible similar dynamic in the soil formation and post depositional processes. From the results reported by Qi et al. (2017), ROC fraction included condensed pyrogenic carbon, condensed aromatic humic carbon and other non-reactive organic carbon fractions. The ROC can be related to the pyrogenic products that can, in turn, contribute to an increase of the more labile TOC in burial conditions. Therefore, in the context of DE, TOC input may partially derive from the significant contribution of charcoal and biomass burning-derived particles, as suggested by the significant positive correlation coefficients between TOC and CHAR500 (r = 0.53, p-value <0.01), CHAR250 (r = 0.50, p-value <0.01) and CHAR125 (r = 0.52, p-value <0.01). The pyrogenic source of ROC can be also supported by the correlation coefficient of ROC vs CHAR500 (r = 0.57, p-value <0.01), CHAR250 (r = 0.55, p-value <0.01), CHAR125 (r = 0.61, p-value <0.01) and CHAR63 (r = 0.38, p-value <0.05).

Table 3.10. Pearson correlation coefficient (significance expressed as * when p-value ≤ 0.05, ** when p-value ≤ 0.01) between organic (TOC), recalcitrant (ROC) and inorganic carbon (TIC), dimensional charcoal counts (1000,500,250,125 and 63µm) and Raman parameters (Band G, Band D, I_D/I_G and I_V/I_G).

	ROC	TIC	CHAR1000	CHAR500	CHAR250	CHAR125	CHAR63	Band G (cm^{-1})	Band D (cm^{-1})	I_D / I_G	I_V / I_G	L_a (nm)
TOC	0.60 **	-0.72 **	0.28	0.53 **	0.50 **	0.52 **	0.21	-0.32	0.24	0.20	-0.16	-0.14
ROC	1	-0.65 **	0.36	0.57 **	0.55 **	0.61 **	0.38 *	-0.24	0.39	0.17	-0.02	-0.13
TIC		1	-0.04	-0.39 **	-0.25	-0.29	-0.20	-0.17	0.35	-0.30	-0.08	0.30
CHAR (1000)			1	0.80 **	0.78 **	0.84 **	0.55 **	-0.31	0.25	-0.04	-0.10	0.18
CHAR (500)				1	0.81 **	0.78 **	0.41 **	-0.35	0.22	0.34	-0.10	-0.22
CHAR (250)					1	0.92 **	0.62 **	-0.29	0.19	-0.01	-0.15	0.11
CHAR (125)						1	0.72 **	-0.37	0.26	0.03	-0.44	0.07
CHAR (63)							1	-0.42	0.28	-0.23	0.34	0.26
Band G (cm^{-1})								1	-0.09	-0.24	-0.52 **	0.17
Band D (cm^{-1})									1	-0.11	0.30	0.11
I_D / I_G										1	0.11	-
I_V / I_G											1	-0.13

Such evidence is in agreement with the researches of Zethof et al., (2019) that reported ROC as a valuable tool to measure the content of graphitic carbon originated from biomass burning products.

When TOC and ROC are compared to TIC, they display a significant negative correlation ($r = -0.72$, p-value <0.01; $r = -0.65$, p-value <0.01, respectively). The behavior of organic and inorganic fractions in soils and sediments is often hardly comprehensible: positive and negative correlations were detected also in similar environmental conditions (Leogrande et al., 2021). Indeed, when the investigated site is a closed system (e.g. a pond or a lake without inflows/outflows), the endogenic input of both TIC and TOC are positively and significantly correlated (Ju et al., 2010). Conversely, in an open system, such as DE deposits, different inputs of TIC and TOC that derive from human activities may have affected the soil conditions. Leogrande et al. (2021) found negative, although not always significant, correlation between TOC and TIC in farms and orchard's soils in the Mediterranean area. The different oxidizing conditions induced by a relatively high concentration of carbonates could have promoted the increase of TIC at the expense of TOC. Other studies proposed that the negative correlation between TOC and TIC can be related to the degradation of organic matter by water dissolution or availability of microorganism able to mineralize the organic matter (Lu et al., 2020). In DE, the significant negative correlation could be influenced also by the effect of the Ca aggregates, able to stabilize TOC, while CO_2 production can cause a pH decrease, and in turn increase the dissolution of carbonate (Sartori et al., 2007; Shi et al., 2017). The observed negative correlation between TIC and TOC could be due to the continuous input of biomass and organic material to the soil, which dilutes the original matrix. The continuous presence of human activities over centuries in the studied sites could have promoted the input of external material and its mixing with the original matrix, forcing the system to rearrange and re-equilibrate.

The distribution of charcoal particles in the different dimensional size classes are all significant and positively correlated, indicating that such fragments might have the same source. Their exact origin is difficult to interpret because of the multitude of processes that could alter charcoal characteristics (e.g., bioturbation, erosion and freezing) (Conedera et al., 2009a). A very local source of the charred particles is suggested by the presence of >125 μm charcoal (Conedera et al., 2009b;Whitlock and Larsen, 2002). However, it cannot be excluded that the significant positive correlation

found between the number of charcoal particles per size class could be ascribed to the fragmentation of larger charcoal fragments due to handling processes of soil.

The significant negative correlation between I_V/I_G and the G band position (r = -0.52, p-value <0.01) can be likely associated to graphitization processes, as the wavenumber of the G band shifts toward higher values while I_G increases in intensity (I_V/I_G decreases) when graphitic structures prevails over amorphous ones (Inoue et al., 2017; McDonald-Wharry et al., 2013).

3.7 Soil morphology and microscopic features: summary

The two medieval sites (VP and SPM) belong to the same stratigraphic level and have a close location, indeed ^{14}C dates indicate that they encompass the same period. However, radiocarbon ages did not provide a chronological sequence, suggesting that the matrix underwent possible mixing phenomena, as also shown by the thin section observation. Indeed, independently on the fraction of carbon used for the radiocarbon analysis (i.e., charcoal or humin), the ages do not increase with depth. Nevertheless, the slight differences obtained when dating humin or charcoal suggest a possible initial anthropization period, resulting in an enrichment in organic matter (1st and 2nd century AD) before the cultural layer (4th – early 6th century AD) was deposited. Humin carbon represents a part of the matrix originally deposited (van der Plicht et al., 2019; van Mourik et al., 2012), enriched in coarser charcoal fragments during the following centuries.

The dumping of heterogeneous material, where older sedimentary matrices and older components are mixed with inclusions dating closer to the time of formation, seems the most probable explanation for the DE formation at SPM. Indeed, the presence of hiatuses, and few signs of bioturbation suggest a fast-growing DE in both SPM and VP. The more pronounced bioturbation process observed in CDL indicates a slower accumulation of the soil.

The presence of metal slags and droplets produced during ironworking (Angelini et al., 2017) in SPM and especially in VP, suggests the existence of metallurgical production

activities, also confirmed by the analysis of trace elements. The differences in dimensions and abundance of the metallic droplets between VP and SPM indicate that VP may have been a localized site for metal smoldering, and SPM has been only indirectly affected by this activity.

High P content at all sites points to a relevant anthropogenic influence (Patten et al., 2012), since it can be associated to excrement and byre addition, and/or to the dumping of biomass burning residues (Devos, 2018). This activity is the most probable cause of P excess in VP, in agreement with the more abundant presence of charred fragments and ROC in this site with respect to SPM and CDL.

From the distribution of charcoal fragments, trace elements analysis, and organic/inorganic carbon determination it is possible to highlight that in both SPM and VP no trends or internal differentiations were observed. This suggests that several processes of soil exploitation (backfilling and dumping of material mixed with the matrix and sometimes reworked), occurred simultaneously or in a limited time period, in agreement with the occurrence of different processes of soil exploitation and with the absence of chronological distribution of the radiocarbon dates.

The significant differences in terms of charcoal fractions between SPM and VP, suggest that the horizontal variability in DE, likely related to different practices such as agriculture or waste management, is relevant even for short-range distances. VP and SPM also show a different fingerprint when considering TOC, ROC and charcoal distributions, although these two sites belong to DE deposits that can be correlated. The different charcoal distribution may be associated to distinct dynamics of production and/or transport of the biomass burning residues. This might in turn be related to independent anthropogenic activities that occurred in the specific site (Davidson et al., 2006), such as the metallurgical activities described above.

Finally, Raman results obtained on charcoal particles showed that fragments from CDL are characterized by a higher level of structural disorder with respect to the charred particles selected for the SPM and VP sites. This could be explained by the variation in age, as CDL is likely older than the others. Being older and possibly more degraded, charcoal had more disordered structures with a lower fraction of graphitized carbon (Smidt et al., 2020, 2017). This aging process may be ascribed to environmental factors, such as pH and humidity of the soil, microbial activities or water erosion (Arroyo-Kalin, 2020; Inoue et al., 2017; Rumpel et al., 2015). The chemical and physical changes of charred particles due to interactions with the soil were studied by Cheng et al., (2008) and Inoue et al. (2017) and it was suggested that the main reason for the

changes is physical and biological erosion. Additionally, higher values of I_V/I_G in CDL may be indicative of lower burning temperatures. This implies that the more ancient charcoal of CDL was produced at lower carbonization temperatures, providing a limited removal of amorphous structures and a scarce formation of graphene-like networks (McDonald-Wharry et al., 2013). On the other hand, the similarity between SPM and VP supports the assumption that SPM and VP are coeval, while higher carbonization temperatures may be related to high-energy demand processes, such as smelting (as previously observed) or ceramic manufacturing.

4 Plant and biomass burning molecular proxies

The analysis of different chemical biomarkers from sedimentary deposits can assist in the reconstruction of past environmental conditions. Within these chemical biomarkers, the analysis of fire proxies is a valuable approach for the determination of past human presence. Nonetheless, plant biomarkers are useful tools to infer the environmental conditions where, or because of which, anthropization occurred. In this section, the results and the discussion regarding environmental indicators (linear alkanes and miliacin) and fire proxies (polycyclic aromatic hydrocarbons and anhydrous sugars) are reported.

The whole set of molecular proxies was determined in the samples taken from the three sites via Pigna (VP), vicolo San Pietro in Monastero (SPM), and Castelar di Leppia (CDL), located in the province of Verona. For the two sites VP and SPM the samples were collected following two vertical parallel profiles, as reported in Figure 2.1. Table 4.1 reports the number of soil samples analyzed for each profile and site.

Table 4.1. Number of samples analyzed from each profile and site.

Profile/site	Number of samples
SPM I	6
SPM II	12
SPM	18
VP I	6
VP II	6
VP	12
CDL	3

4.1 Linear alkanes and miliacin

The concentration of 26 linear alkanes (from decane $C_{10}H_{22}$ to pentatriacontane $C_{35}H_{72}$) was determined together with miliacin in the DE samples from the sites of via Pigna (VP), vicolo San Pietro in Monastero (SPM), and Castelar di Leppia (CDL), located in

the province of Verona. Please refer to table 4.1 for the number of soil samples analyzed for each profile and site.

The approach used for the investigation of fire proxies, microscopic particles, and carbon fractions (Chapter 3) has been applied to the plant biomarkers: *n*-alkanes and miliacin, because also in this case it was impossible to identify trends or pattern following the depth profiles. The profiles of concentration of all the linear alkanes and miliacin for each sample are reported in Appendix D.

The concentration of *n*-alkanes spans between 3 ng and 3.5 μg per gram of soil extracted. Figure 4.1 describes the distributions of *n*-alkanes for profiles and sites, where the bar represents the mean value and the error bar represents the 95% confidence interval. From Figure 4.1 it is possible to observe slightly different distributions between sites, with SPM presenting higher *n*-alkanes concentrations, whereas alkanes with shorter chains are absent or rarely present in VP and CDL. In SPM the most abundant congeners are triacontane (C_{30}) (1000 ± 500 ng g^{-1}), dotriacontane (C_{32}) (900 ± 500 ng g^{-1}), and tetratriacontane (C_{34}) (600 ± 300 ng g^{-1}). The same *n*-alkanes are those mainly present in the profile SPM I, with concentrations of 1200 ± 1000 ng g^{-1} of C_{32}, 1000 ± 1000 ng g^{-1} of C_{30} and 850 ± 700 ng g^{-1} of C_{34}; whereas the most abundant linear alkanes in the SPM II profile are C_{30} (900 ± 700 ng g^{-1}), C_{32} (750 ± 600 ng g^{-1}) and pentatriacontane (C_{35}) (720 ng g^{-1}).

Similarly, VP, VP I and VP II, are enriched in C_{32}, C_{30}, and C_{28}. In particular, concentrations of 700 ± 400 ng g^{-1}, 600 ± 500 ng g^{-1}, and 350 ± 250 ng g^{-1} were found for VP. In VP I and VP II the same congeners account respectively for: C_{32} 730 ± 800 ng g^{-1} and 700 ± 1000 ng g^{-1}, C_{30} 600 ± 800 ng g^{-1} and 600 ± 1700 ng g^{-1}, and C_{28} 360 ± 200 ng g^{-1} and 340 ± 400 ng g^{-1}.

Conversely, in CDL the main contribution is given by C_{35} (1100 ng g^{-1}), C_{22} (430 ng g^{-1}), and C_{31} (400 ± 400 ng g^{-1}).

Miliacin, instead, is mainly present in CDL (220 ± 200 ng g^{-1}), followed by SPM (150 ± 30 ng g^{-1}), and VP (90 ± 10 ng g^{-1}).

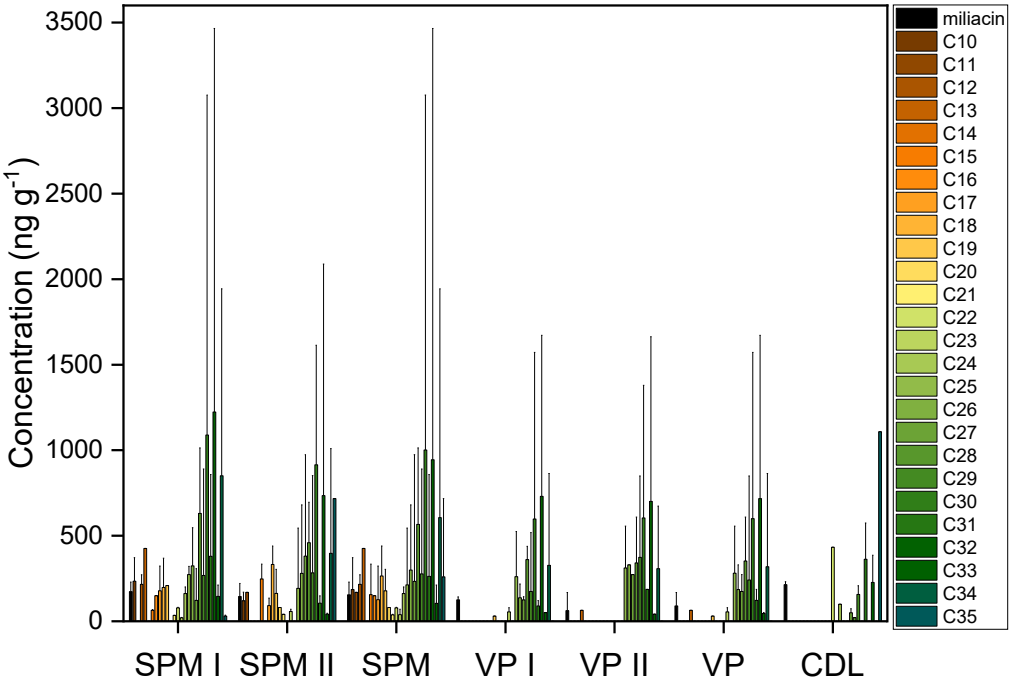

Figure 4.1. Distribution of *n*-alkanes for profiles (SPM I, SPM II, VP I, and VP II) and sites (VP, SPM, CDL) bar heights represent the mean values and the error bars correspond to the confidence interval (95%).

For all the sites, the higher content of linear alkanes containing more than 25 carbon atoms can be noticed, while the concentration of *n*-alkanes with less than 25 carbon atoms was found to be lower than the LOD in several samples. Indeed, even if the concentration of single alkanes is aligned with values reported in the literature (Xie et al., 2003; Zech et al., 2013), the presence of congeners below the LOD makes it difficult to calculate diagnostic ratios and make comparison with published results. For this reason, Table 4.2 reports the alkanes distribution grouped according to the possible main sources described in the literature (Section 1.2.3): short chain (C_{11}-C_{20}) derived from bacteria, plankton, and aquatic microorganisms (Cranwell et al., 1987; Pisani et al., 2013), medium chain length (C_{20}-C_{25}) indicative of submerged and floating macrophytes (Cranwell et al., 1987; Ficken et al., 2000), and long-chain linear alkanes of terrestrial vegetation of trees and shrubs (C_{26}-C_{29}) and grasses (C_{30}-C_{35}) (Ficken et al., 2000; Killops and Killops, 2005; Zech et al., 2013, 2010). In addition, Table 4.2 reports the average chain length (ACL), as a reliable indicator of the most predominant alkanes for each profile and site, while the odd-even predominance (OEP) ratio is known as useful tool to evaluate the degree of preservation of the original plant input (Zech et al., 2010); lastly, the concentration of miliacin is reported. Table 4.2 accounts for the mean values and the confidence intervals for each site and profile, for the vegetal

variables described, in brackets the number of observations considered for the calculation (outliers excluded).

Table 4.2. Mean values and confidence intervals for *n*-alkanes proxies and miliacin calculated for profiles and sites. In brackets the number of observations for each distribution are reported.

	C_{11}-C_{19}	C_{20}-C_{25}	C_{26}-C_{29}	C_{30}-C_{35}	ACL	OEP	Miliacin
SPMI	750 ± 650 (4)	300 ± 300 (4)	1500 ± 1000 (4)	3000 ± 1000 (4)	28±2 (4)	0.2±0.1 (4)	170 ± 60 (5)
SPMII	600 ± 500 (4)	200 ± 150 (7)	600 ± 600 (9)	3000 ± 2000 (5)	28±1 (11)	0.15 ± 0.05 (5)	140 ± 40 (9)
SPM	600 ± 300 (8)	200 ± 100 (11)	900 ± 500 (13)	3000 ± 1000 (9)	28±1 (15)	0.2±0.1 (9)	150 ± 30 (14)
VPI	30 (1)	300 ± 300 (3)	700 ± 600 (4)	1700 ± 2000 (2)	29 ± 3 (6)	0.2 ± 0.1 (5)	120 ± 20 (4)
VPII	60 (2)	300 ± 350 (2)	600 ± 900 (4)	1500 ± 2000 (2)	29 ± 3 (5)	0.3 ± 0.2 (3)	60 ± 80 (5)
VP	50 ± 20 (3)	300 ± 300 (5)	600 ± 600 (8)	1600 ± 1200 (4)	29 ± 1 (11)	0.2 ± 0.1 (8)	90 ± 10 (9)
CDL	-	500 (1)	200 ± 20 (2)	1000 ± 600 (3)	31 ± 1 (3)	20 (1)	220 ± 200 (2)

The low-weight *n*-alkanes (C_{11}-C_{19}) are absent in CDL, and present variable concentrations in SPM (600 ± 300 ng g^{-1}) and VP (50 ± 20 ng g^{-1}). For the *n*-alkanes between C_{20} and C_{25}, SPM presents the lowest content (200 ± 300 ng g^{-1}) if compared to the other sites, which display slightly higher but comparable values (VP: 300 ± 300 ng g^{-1}, CDL: 500 ng g^{-1}). The input of trees (C_{26}-C_{29}) is again more relevant in SPM (1500 ± 1000 ng g^{-1}) than in VP (600 ± 600 ng g^{-1}) and CDL (200 ± 20 ng g^{-1}). Lastly, the *n*-alkanes associated to the presence of grasses and shrubs (C_{30}-C_{35}) are the most prominent contribution to all the sites, following a decreasing concentration: SPM (3000 ± 1000 ng g^{-1}) > VP (1600± 1200 ng g^{-1}) > CDL (1000±600 ng g^{-1}). The predominance of the terrestrial vegetation in all the sites is also evidenced by the values of the ACL: for CDL it is 31 ± 1, in VP 29 ± 1, and SPM 28 ± 1.

For VP and SPM, the OEP ratio presents values <1, meaning a strong presence of *n*-alkanes with an even number of carbon atoms. On the other hand, CDL has just one value (21) because in the other two samples the *n*-alkanes with an even number of carbon atoms were below the detection limit, meaning an exclusivity of chains with odd number of carbon atoms. These differences suggest that it is possible that distinct degradation process of the organic matter have been operated by soil microbiota.

Table 4.3 reports the p-values of the comparison between sites and profiles. The statistical investigation between profiles and sites consisted of Welch's two Sample t-test for normal distributions and the Wilcoxon-Mann-Whitney rank test for the

distributions where normality was not achieved. Normality was tested with the Shapiro-Wilk normality test (α=0.05).

Table 4.3. The comparison of the profiles within the same sites do not show significant differences for both SPM and VP, while p-values below α=0.05 are underlined

	C_{11}-C_{19}	C_{20}-C_{25}	C_{26}-C_{29}	C_{30}-C_{35}	ACL	OEP	Miliacin
SPMI vs SPMII	0.5	0.5	0.1	0.8	0.4	0.8	0.4
VPI vs VPII	1	1	0.4	0.9	0.9	0.3	0.1
VP vs SPM	<u>0.002</u>	0.6	0.4	0.06	0.7	0.9	<u>0.04</u>
VP vs CDL	-	0.1	0.1	0.9	<u>0.02</u>	<u>$<10^{-16}$</u>	<u>0.0002</u>
SPM vs CDL	-	<u>$2 \cdot 10^{-4}$</u>	0.2	<u>0.008</u>	<u>0.0003</u>	<u>$<10^{-16}$</u>	<u>0.001</u>

The p-values resulting from the comparison between the profiles indicate that there is no evidence of significant differences within the same site. For the group of alkanes C_{11}-C_{19}, VP and SPM are significantly different. As reported previously, SPM presents higher concentrations of all the alkanes, including the light congeners. For all the other groups of linear alkanes, there are no other significant differences between the two medieval sites VP and SPM. Some important differences in the distribution of alkanes can be noticed from the comparison between SPM and CDL: the content of the *n*-alkanes C_{20}-C_{25} is significantly higher in CDL than in SPM (p-value = $2 \cdot 10^{-4}$), while SPM is enriched in long-chain alkanes C_{30}-C_{35} (p-value = 0.008).

Significant differences are observed when sites are compared in terms of ACL values: CDL vs both VP and SPM present p-values below 0.05. ACL values are similar for VP and SPM and centered around 28-29, whereas it is higher for CDL, with a mean value of 31. These differences suggest that the type of terrestrial plants present in the sites was not alike: in CDL there could have been a predominance of grasses and shrubs instead of tall trees (Eglinton et al., 1962; Fang et al., 2014; Ficken et al., 2000; Silva et al., 2012; Zech et al., 2013, 2010). The relatively higher presence of grasses input could possibly be supported by the greater concentration of miliacin detected in CDL in comparison to both SPM and VP, possibly indicating a diverse use of the for agricultural purposes, or a different land cover. The content of miliacin is also significantly different between SPM and VP, suggesting a differentiated use of the soil even in a short-range distance, exploited for productive purposes in SPM.

The p-values resulting from the comparison of OEP between sites highlight again the difference between the medieval sites (SPM and VP) and the Bronze Age layer (CDL),

with values much lower than 0.05 ($<10^{-16}$). The absence of odd over even carbon predominance, (i.e., OEP<1), can be indicative of a biodegradation mechanism (Grimalt et al., 1988). Indeed, CDL seems to be less affected by the biodegradation, showing a higher odd carbon presence with respect to SPM and VP. As reported above, the DE layer from CDL accounts for a lower concentration of *n*-alkanes with even number of carbon atoms than SPM and VP although it is older. This could be due to different soil conditions in the two areas: linear alkanes are highly reduced molecules, that, to be degraded, require the presence of an electron acceptors, mainly represented by oxygen (Chandra et al., 2013). VP and SPM present OEP values around 0.2, indicating higher conversion of the alkanes. This could be favored by well oxygenated soil, due to mixing processes, as suggested by the absence of a precise correspondence between chronology and depth in the analyzed profiles (Section 3.1). Conversely, the higher OEP in CDL could be influenced by a progressive decreasing of oxygen in soil, because of a depleted soil without signs of backfilling, even if the presence of earthworm's channels suggests the presence of oxygen in the soil. Alternatively, it has been reported that other electron acceptors such as nitrate, iron, bicarbonate, nitrous oxide and sulphate, can increase the rate of conversion of linear hydrocarbons (Chandra et al., 2013; Moucawi et al., 1981). Hence, the enrichment of VP in Fe (Table 3.3, section 3.3) might have promoted the biodegradation rate of the alkanes.

Many studies focusing on the capability of fungi and bacteria to degrade the saturated and unsaturated hydrocarbons listed the hydrocarbon-utilizing genres and species in different soil and sediments conditions (Chandra et al., 2013; Daccò et al., 2020, and references therein). However, the dearth of knowledge on the microfauna populations that colonize DE and cultural layers in European context makes it difficult to outline a specific type of degradation process affecting the samples of this study.

4.1.1 Plant proxy: summary

As discussed in these paragraphs, the distribution of the linear alkanes can be associated to a differentiated plant input and, in this framework, *n*-alkanes can be used for the reconstruction of the paleoenvironment and the changes in land coverage through time. At the same time, the relative distribution depends also on the degradation degree of the organic matter. As reported, in all the investigated sites the main contribution of alkanes seems to come from terrestrial plants input, as suggested by the strong abundance of heavier *n*-alkanes and the high ACL values. The predominance of terrestrial plants is also supported by the concentrations of miliacin detected in the three sites. The degree of degradation of the hydrocarbons, expressed by the OEP suggested different soil conditions: VP and SPM present a more expressed degree of conversion, possibly due to a periodic oxygenation attributable to a remixing of the soil, whereas CDL seem to be less affected by biodegradation phenomena, possibly indicating a lower level of soil manipulation.

4.2 Polycyclic aromatic hydrocarbons (PAHs)

As what previously discussed for TOC, ROC and TIC (Section 3.4) and *n*-alkanes distribution (Section 4.1), also in the case of polycyclic aromatic hydrocarbons (PAHs) there was no evidence of vertical trends along the profiles. For this reason, results are reported here as mean values for profile section/site, while Appendix D lists the PAH concentrations for each sample.

The overall concentrations of PAHs spanned between a minimum of 1.2 pg g^{-1} and a maximum of 68 ng g^{-1} (referred to the dry-weighted mass of soil extracted). These concentrations are slightly lower, but aligned to the ones found in other studies where archaeological soils were investigated (Tan et al., 2020; Zhang et al., 2013).

Figure 4.2 shows the distribution of PAHs for each site/profile investigated. Bars represent mean values, and the error bars correspond to the confidence interval (95%)

of the distribution. In general, higher PAH concentrations were observed in SPM, followed by VP and CDL.

As shown in Figure 4.2, the most abundant PAH in SPM is phenanthrene (Phe), followed by retene (Ret), and benzo(e)pyrene (BeP), with the concentrations (mean ± 95% confidence interval) of 11 ± 5 ng g^{-1}, 7 ± 2 ng g^{-1}, and 6 ± 4 ng g^{-1}, respectively. In the two profiles SPM I and SPM II the three most abundant congeners were respectively: BeP (10 ± 12 ng g^{-1}), indeno[1, 2, 3,-c, d]pyrene (IP) (7 ± 9 ng g^{-1}) for SPM I, and fluoranthene (Fla) (6 ± 9 ng g^{-1}), and Phe (14 ± 6 ng g^{-1}), retene (Ret) (8 ± 2 ng g^{-1}), BeP (4 ± 4 ng g^{-1}) for SPM II.

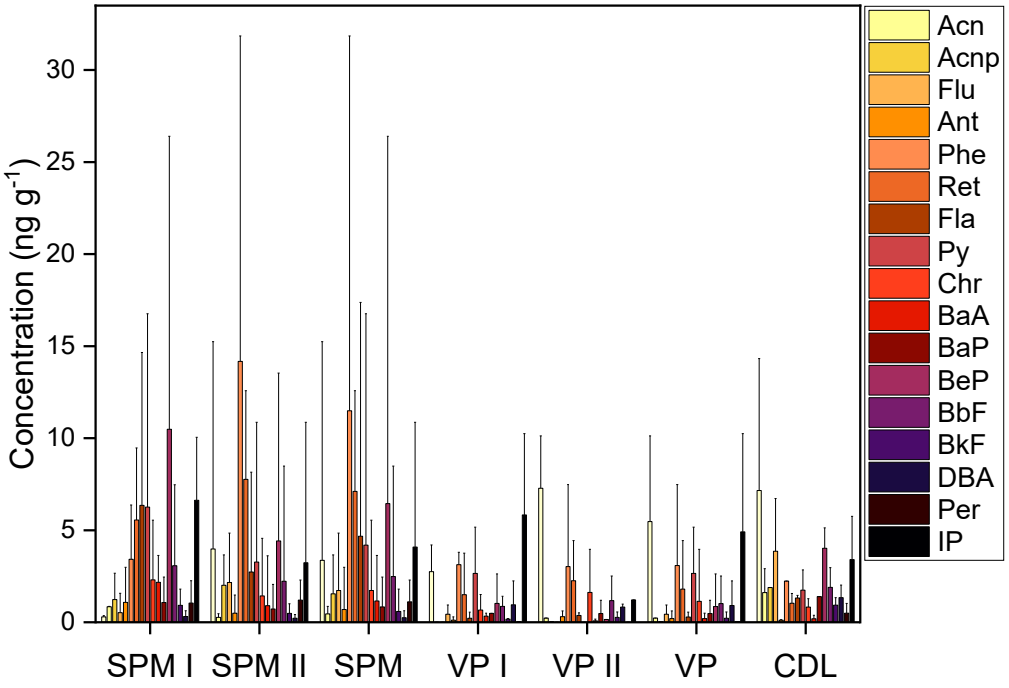

Figure 4.2. PAH distribution for profiles (SPM I, SPM II, VP I, and VP II) and sites (VP, SPM, CDL) bars represent the mean values, and the error bars correspond to the confidence interval (95%).

In VP the most abundant PAH is naphthalene (Nap), followed by IP and Phe with concentrations of 6 ± 5 ng g^{-1}, 5 ± 5 ng g^{-1}, and 3 ± 1 ng g^{-1}. For the two profiles VP I and VP II, the most abundant congeners are Phe (3 ± 1 ng g^{-1}), Nap (3 ± 2 ng g^{-1}), IP (6 ± 7 ng g^{-1}) for the first profile, whereas for VP II Nap (7 ± 10 ng g^{-1}), Phe (3 ± 5 ng g^{-1}), Ret (2 ± 2 ng g^{-1}) are the most present PAHs.

Finally, the first three PAHs detected in CDL are Nap, benzo(a)pyrene (BaP), and fluorene (Flu) with concentrations, of 7 ± 17 ng g^{-1}, 5 ± 14 ng g^{-1}, and 4 ± 7 ng g^{-1}, respectively.

Table 4.4 reports the mean and confidence intervals for the sum of PAHs as concentration (ng g^{-1}), and as percentage relative to the total sum of PAHs. The 2 and 3-ring congeners (Nap, Acn, Acnp, Flu, Ant, Phe, Ret) were grouped together as low-weight PAHs (hereafter LW PAH), PAHs formed by 4 condensed aromatic rings (Fla, Py, Chr, BaA) were classified as medium weight PAHs (hereafter MW PAH), while the group of high-weight PAH (hereafter HW PAH) was formed by the 5 and 6 rings PAHs (BaP, BeP, BbF, BkF, DBA, Per, IP, BghiP). The number of samples used to calculate mean values and confidence intervals is reported in brackets. Before calculating the parameters for each distribution, outliers were deleted following the Dixon test for one outlier (α=0.05), and the Grubbs test for one tail two outliers (α=0.05).

Table 4.4. Mean values and confidence intervals of PAH groups for profiles and sites at vicolo San Pietro in Monastero (SPM I, SPM II, and SPM) via Pigna (VP I, VP II, and VP), and Castelar di Leppia (CDL).

	LW PAH (ng g^{-1})	**MW PAH (ng g^{-1})**	**HW PAH (ng g^{-1})**	**LW PAH %**	**MW PAH %**	**HW PAH %**
SPMI	9 ± 5 (6)	7 ± 10 (5)	20 ± 18 (6)	40 ± 30 (6)	10 ± 6 (5)	30 ± 20 (6)
SPMII	28 ± 10 (12)	7 ± 5 (12)	12 ± 7 (12)	40 ± 20 (12)	11 ± 5 (12)	20 ± 10 (12)
SPM	21 ± 8 (18)	7 ± 4 (17)	15 ± 7 (18)	50 ± 10 (18)	11 ± 4 (17)	24 ± 8 (18)
VPI	6 ± 4 (6)	2 ± 3 (5)	10 ± 9 (6)	20 ± 10 (6)	12 ± 10 (5)	30 ± 10 (6)
VPII	7 ± 7 (6)	2 ± 2 (4)	4 ± 4 (6)	60 ± 30 (6)	13 ± 10 (4)	20 ± 10 (6)
VP	7 ± 4 (12)	2 ± 3 (9)	7 ± 9 (12)	40 ± 20 (12)	13 ± 10 (9)	20 ± 10 (12)
CDL	15 ± 30 (3)	4 ± 4 (2)	10 ± 10 (3)	50 ± 20 (3)	10 ± 20 (2)	20 ± 50 (3)

Lighter PAHs typically dominate the distribution, followed by heavier congeners, whereas medium weight congeners are the least present. Indeed, the sum of LW PAH constitutes 40-60% of the total, 4-ring congeners are attested to around 10% of the total PAHs, and heavier congeners are around 20-30% of the PAHs for all the sites.

The distributions of each class of PAHs for the profiles within the same site and between different sites were compared using the Welch two Sample t-test for normal distributions and the Wilcoxon-Mann-Whitney rank test for distributions where normality was not achieved. Normality was tested with the Shapiro-Wilk normality test (α=0.05). The p-values of the comparison test are reported in Table 4.5. The p-values < α (0.05) are reported underlined, meaning that the two distributions compared for the specific biomarker are significantly different for α=0.05.

Table 4.5. Comparison between mean values of PAH groups. p-values obtained using the t-Student or the Mann-Whitney test depending on the normality of the distribution previously checked with the Shapiro-Wilk test. p-values < 0.05 (underlined) indicates a significant difference between the mean values.

	LW PAH	MW PAH	HW PAH
SPMI vs SPMII	<u>0.001</u>	0.7	0.3
VPI vs VPII	0.6	0.7	0.2
VP vs SPM	<u>0.002</u>	<u>0.03</u>	0.09
VP vs CDL	0.4	<u>0.001</u>	1
SPM vs CDL	0.4	0.1	0.6

Generally, the intra-site comparisons (profile I vs. profile II) show non-significant p-values for the sum of PAHs (Table 4.5), this means that the two profiles do not show differences, similarly to the comparisons shown in Chapter 3. The only p-value below 0.05 resulting from the intra-site comparison is the one between the profiles SPM I and SPM II for the concentration of LW PAH. In fact, the content of LW PAH for the profile SPM II (28 ± 10 ng g^{-1}) is higher than SPM I (9 ± 5 ng g^{-1}) (Table 4.4), probably because of the higher concentration of Phe in SPM II with respect to SPM I. This is consistent with the occurrence of signs of burned soil exclusively in the profile SPM II, as determined during the micromorphological observation (Appendix B).

Significant differences also arise when comparing the LW and MW PAH contents of VP and SPM, with p-values of 0.002 and 0.03 respectively (Table 4.5). LW PAH in VP and SPM accounted for 7 ± 4 ng g^{-1} and 21 ± 8 ng g^{-1}, respectively, whereas MW PAH concentrations are 2 ± 3 ng g^{-1} and 7 ± 4 ng g^{-1}. In general, higher concentrations of PAHs were determined in SPM with respect to the other two sites (Figure 4.2). From the evaluation between the PAH content determined in SPM and CDL, none of the p-values is below the level of significance, meaning that even if CDL presents lower concentrations for each PAH group (15 ± 30 ng g^{-1} for LW PAH, 4 ± 4 ng g^{-1} for MW PAH, and 10 ± 10 ng g^{-1} for HW PAH), these differences are not significant when compared to SPM. CDL displays higher concentrations of all the PAHs than VP (7 ± 4 ng g^{-1} for LW PAH, 2 ± 3 ng g^{-1} for MW PAH, and 7 ± 9 ng g^{-1} for HW PAH, table 4.1), nevertheless, the only significant difference is related to MW PAH with a p-value of 0.001 (Table 4.5).

Globally, the studied sites can be classified by the content of PAH in the following order SPM > CDL > VP.

Besides the classification into different weight congeners, PAHs are also commonly classified following their main source of production: pyrogenic or petrogenic. The classification into pyrogenic PAH was done following Zou et al., (2010) and D'Anjou et al., (2012): they considered as indicative of pyrolytic processes the congeners Fla, Py, BaP, BeP, BbF, BkF, BghiP. On the other hand, petrogenic PAHs were classified following Wang et al., (1998) and Yunker et al., (2002), identifying Nap, Flu, Py, Chr, and BghiP, as petrogenic PAHs. However, the interpretation of some of these analytes can be difficult and/or ambiguous since pyrene and benzo(g,h,i)perylene are associated to both classes. Other studies report the HW PAHs (formed by four-six condensed rings) as of pyrogenic origin, and associate the abundance of one to 3-ring congeners to a petrogenic origin (Wilcke, 2007; Zhang et al., 2013). The origin of PAHs in DE samples was also investigated using ratios between specific congeners as reported in section 1.2.2.

Tables 4.6 and 4.7 summarize the results of the calculation and comparison between proxies for sites and profiles. Specifically, the mean values and the confidence intervals of groups of biomarkers for profiles and sites are reported in Table 4.6, while Table 4.7 shows the p-values resulting from the comparison between profiles and sites.

Table 4.6. Mean values and confidence intervals of PAH ratios and pyrogenic and petrogenic PAHs for profiles and sites at vicolo San Pietro in Monastero (SPM I, SPM II, and SPM) via Pigna (VP I, VP II, and VP), and Castelar di Leppia (CDL).

	Pyrogenic PAH	Petrogenic PAH	IP/(IP+BghiP)	Ret/(Ret+Chr)
SPMI	25 ± 24 (6)	10 ± 10 (6)	0.5 ± 0.2 (2)	0.8 ± 0.2 (4)
SPMII	14 ± 9 (12)	12 ± 6 (12)	0.5 ± 0.1 (7)	0.8 ± 0.1 (11)
SPM	17 ± 9 (18)	11 ± 5 (18)	0.5 ± 0.1 (9)	0.8 ± 0.1 (15)
VPI	6 ± 5 (6)	7 ± 6 (5)	0.8 (1)	0.8 ± 0.2 (3)
VPII	4 ± 4 (6)	8 ± 10 (5)	0.8 (1)	0.9 ± 0.1 (3)
VP	5 ± 5 (12)	8 ± 6 (10)	0.84 ± 0.02 (2)	0.8 ± 0.1 (6)
CDL	12 ± 9 (2)	14 ± 30 (3)	0.9 (1)	0.5 (1)

Table 4.7. Comparison between mean values of PAH groups. p-values obtained using the t-Student or the Mann-Whitney test depending on the normality of the distribution previously checked with the Shapiro-Wilk test. p-values < 0.05 (underlined) indicate a significant difference between the means.

	Pyrogenic PAH	Petrogenic PAH	IP/(IP+BghiP)	Ret/(Ret+Chr)
SPM I vs SPM II	0.6	1	0.8	0.2
VP I vs VP II	0.3	0.9	1	0.08
VP vs SPM	0.02	0.3	0.03	0.5
VP vs CDL	0.0002	0.5	0.6	0.002
SPM vs CDL	0.2	0.8	0.00005	0.0006

The amounts of both pyrogenic and petrogenic PAHs are higher in SPM and CDL than in VP, with pyrogenic PAHs concentrations of 17 ± 9 ng g^{-1}, 12 ± 9 ng g^{-1}, and 8 ± 6 ng g^{-1} respectively, and petrogenic PAHs of 11 ± 5 ng g^{-1}, 14 ± 30 ng g^{-1}, and 8 ± 6 ng g^{-1}.

Consistently with the comparisons shown in Table 4.4, the content of pyrogenic PAHs is significantly different for VP (5 ± 5 ng g^{-1}) when compared to both SPM (17 ± 9 ng g^{-1}, p-value = 0.02) and CDL (12 ± 9 ng g^{-1}, p-value = 0.0002), whereas the content of pyrogenic PAHs is similar between CDL and SPM. In both sites, in fact, the sum of pyrogenic congeners shows a higher mean value mostly because of the contribution of the previously discussed LW PAH.

The ratios for the investigation of fire events present some differences when sites are compared: IP/(IP+BghiP), which is indicative of the combustible's fire event (Yunker et al., 2002), displays significant differences between SPM (0.5 ± 0.1) and both VP (0.8, p-value = 0.03) and CDL (0.9, p-value = 0.00005). The value of this ratio has been associated to the type of PAH source: values above 0.5 are indicative of burned biomasses (Yunker et al., 2002), whereas values below 0.5 can indicate the presence of petrogenic PAHs. In VP and CDL massive presence of PAHs from biomass burning was found, while the value obtained for SPM is closer to the threshold (0.5), making difficult the interpretation in terms of source of PAHs.

A deeper investigation of the biomass burning sources is possible calculating the ratio Ret/(Ret+Chr) for the identification of softwood combustion (values above 0.8). The p-values of the comparison of ratios obtained for the three sites indicate that CDL (0.5) is significantly different from both VP (0.8 ± 0.1, p-value = 0.002) and SPM (0.8 ± 0.1, p-value = 0.0006). These values indicate that while the combustion of softwood can be easily hypothesized for SPM and VP, in CDL different types of biomass were burned.

In conclusion, differences have been found for the three investigated sites when looking both at the content of different classes of weight of PAHs, and at the proxies for the identification of PAH sources. SPM resulted to be the site with higher enrichment of PAH, followed by CDL and VP. The source of PAHs was found to be mostly pyrogenic, and a strong input of Ret was determined in SPM and VP, probably related to softwood combustion.

4.3 Anhydrous sugars

Besides PAHs, also levoglucosan (LEVO) is used as biomass burning marker and it is considered more specific than PAHs because it derives solely by the combustion of cellulose. Table 4.8 reports mean values and confidence intervals (95%) for each site and profile for both LEVO and galactosan (GALA), together with the p-values resulted from the comparison between sites and profiles.

LEVO was determined in concentrations of tens of nanograms per gram of soil (spanning between 2 and 120 ng per gram of soil) and these contents are in accordance with the results published by Gao et al., (2016), the only available publication focusing on archaeological soils to the best of our knowledge. Mannosan (MANNO) was almost absent in the studied samples, only sporadically showing values around 1-2 ng per gram of soil sample. GALA concentration was around 0.1 ng to 14 ng per gram of soil. The two isomers of levoglucosan are reported to be almost always lower than LEVO and this could be related to the different thermal stability of the precursors (hemicelluloses for MANNO and GALA, and cellulose for LEVO) (Callegaro et al., 2018; Kuo et al., 2011; Simoneit, 2002). The concentrations for each sample are reported in the Appendix D.

Table 4.8 Mean values and confidence intervals of Levoglucosan (LEVO) and galactosan (GALA) from two sections at vicolo San Pietro in Monastero (SPM I and SPM II) via Pigna (VP I and VP II), and from Castelar di Leppia (CDL). Comparison between mean values of LEVO and GALA. p-values obtained using the t-Student or the Mann-Whitney test depending on the normality of the distribution previously checked with the Shapiro-Wilk test.

	LEVO (ng g^{-1})	GALA (ng g^{-1})
SPMI	35 ± 20 (5)	3 ± 2 (4)
SPMII	60 ± 30 (10)	7 ± 4 (7)
SPM	50 ± 20 (15)	5 ± 3 (11)
VPI	23 ± 20 (5)	4 ± 3 (3)
VPII	20 ± 14 (5)	3 ± 3 (3)
VP	20 ± 10 (10)	3 ± 2 (6)
CDL	20 ± 50 (2)	-

p-value	LEVO	GALA
SPMI vs SPMII	0.9	0.3
VPI vs VPII	0.2	0.6
VP vs SPM	0.2	0.2
VP vs CDL	0.7	-
SPM vs CDL	0.2	-

Following the same approach presented above, the t-Student or the Mann-Whitney test, depending on the normality of the distribution previously checked with the Shapiro-Wilk test, were employed to identify distributions similarities between sites and profiles. Similar to what was previously observed for the PAH content, in SPM a higher content of LEVO was determined, followed by VP and CDL showing the same mean values but different confidence intervals. Indeed, the large variability of the data within

the same site forbids finding significant inter-sites differences (the p-values are always higher than the significance α=0.05).

4.4 Correlations among biomass burning proxies

This section of the chapter includes the discussion of the whole series of fire proxies presented so far, including those already discussed in Chapter 3. In fact, the purpose of this further dissertation is to outline the possible correlations between variables as well as the similarities and diversities between the studied sites.

Table 4.9 reports the Pearson correlation coefficients of the pairwise correlation between the variables. The significance of the correlation coefficient is specified by the asterisks: * for p-value ≤ 0.05, ** for p-value ≤ 0.01.

The sum of MW PAH is the variable that presents the highest correlation coefficients with respect to all the others. It is positively correlated with LW PAH (0.4*), HW PAH (0.8**), Pyrogenic PAH (0.9**), and negatively with the ratio Ret/(Ret+Chr) (-0.6**). HW PAH is, in turn, correlated with Pyrogenic PAH (0.9**) and anti-correlated with Ret/(Ret+Chr) (-0.7**). Pyrogenic PAH also shows a significant and negative correlation with Ret/(Ret+Chr) (-0.6**). The last significant correlation coefficient in Table 4.6 is the one from the relation between LEVO and his isomer GALA (0.7**).

The existence of significant and positive correlation coefficients between the different groups of PAHs (LW, MW, and HW PAH) could be explained considering the pyrogenic origin as main source for all these proxies (see Section 4.2).

The significant correlation coefficients calculated from the comparison of pyrogenic PAH, MW PAH, and HW PAH can depend on the calculation of these proxies: pyrogenic PAH is the sum of several MW and HW congeners and this could bring to a correlated relation for construction.

Table 4.9. Pearson correlation coefficients (R) between all the fire variables: groups and ratios of PAHs, organic carbon, charred particles, and Raman parameters. Significance is expressed with * if p-value ≤ 0.05, and ** if p-value ≤ 0.01.

	MW PAH	HW PAH	Pyrogenic PAH	IP/(IP+BghiP)	Ret/(Ret+Chr)	LEVO	GALA	TOC	ROC	CHAR 1000	CAHR 500	CHAR 250	CHAR 125	CHAR 63	TOT CHAR	D BAND	G BAND	I_D/I_G	I_V/I_G
LW PAH	0.4*	0.2	0.2	-0.09	0.07	0.2	0.6**	-0.1	-0.3	-0.3	-0.3	-0.2	-0.3	-0.3	-0.3	-0.3	0.1	-0.2	-0.04
MW PAH	1	0.8**	0.9**	-0.04	-0.6**	-0.2	0.2	0-0.1	-0.3	-0.2	-0.2	-0.3	-0.2	-0.3	-0.3	-0.6	-0.05	-0.2	-0.1
HW PAH		1	0.9**	0.1	-0.7**	-0.1	0.3	-0.1	-0.3	-0.2	-0.2	-0.3	-0.2	-0.3	-0.3	-0.2	0.2	0.07	-0.02
Pyrogenic PAH			1	0.2	-0.6**	-0.4	0.005	-0.2	-0.3	-0.2	-0.3	-0.2	-0.2	-0.2	-0.2	-0.07	-0.01	-0.1	0.1
IP/(IP+BghiP)				1	-0.5	-0.2	0.3	0.3	0.3	0.5	0.05	0.4	0.4	0.5	0.6	0.05	-0.5	0.5	0.2
Ret/(Ret+Chr)					1	0.3	-0.3	0.1	0.2	-0.2	-0.04	0.05	-0.1	-0.2	-0.1	0.4	-0.001	0.05	-0.3
LEVO						1	0.7**	-0.2	-0.1	-0.03	-0.08	-0.1	0.003	-0.04	-0.04	-0.01	-0.01	0.4	0.1
GALA							1	-0.2	-0.2	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2	-0.3	0.4	0.3	0.3

The negative and significant correlation of the ratio Ret/(Ret+Chr) with MW PAH, HW PAH, and pyrogenic PAH is less easily interpretable. In fact, the strong input of pyrogenic products derived by softwood combustion should increase the content of LW PAH, in addition to 4, 5, and 6 condensed rings PAHs as minor products (Oros et al., 2006; Oros and Simoneit, 1999). The observed negative correlations, instead, suggest other possible mechanisms involved in the PAH formation. To investigate these possible pathways further, the correlation coefficients of Ret and Chr versus LW PAH, MW PAH, HW PAH, and pyrogenic PAH were calculated and reported in Table 4.10.

Table 4.10. Pearson correlation coefficients between the two PAH Ret and Chr and the PAH groups, * p-value ≤ 0.05, ** p-value ≤ 0.01

	LW PAH	MW PAH	HW PAH	Pyrogenic PAH
Ret	0.7 **	0.3	0.2	0.2
Chr	0.3	0.8**	0.7**	0.6**

Ret is positively and significantly correlated only with the LW congeners, suggesting that the burning of softwood biomasses could possibly mainly increase these congeners.

On the other hand, Chr is positively correlated with all the variables, except LW PAH. It is classified as a MW PAH and it is not present in the calculation of pyrogenic PAHs, however it is strongly correlated both with HW PAH and pyrogenic PAH. Chr has been reported to be one of the most abundant products of hardwood and grasses/straw combustion (Bari et al., 2009; Jenkins et al., 1996), hence, the abundance of chrysene could be related to the presence of fire events where plant residues different from softwood are burned. Such fire events could also promote the production of MW and HW PAH as evidenced by the negative correlation of these congeners with the ratio Ret/(Ret+Chr).

LEVO correlates exclusively with GALA, its isomer. This correlation represents the common origin of both the signal as these are produced by the combustion of biomasses (cellulose and hemicellulose) (Simoneit et al., 1999). The absence of significant correlation coefficients between LEVO (and GALA) and the other fire proxies (both charcoal particles and PAHs) was already mentioned by Gao et al., (2016). The authors explained that the different proxies (e.g., charred particles, PAH, sugars) can indicate different fire production conditions (different temperatures and rates), and/or distinct environmental outcomes. The differences could also be due to different transport and deposition mechanism; in fact, anhydrous sugars are more soluble than PAHs

(Callegaro et al., 2018; Haritash and Kaushik, 2009; Norwood et al., 2013). Moreover, LEVO (and its isomers) could be a less reliable fire proxy in soil samples: it has been demonstrated that some filamentous fungi and yeasts can efficiently metabolize LEVO using the enzyme levoglucosan-kinase. In particular, *Aspergillus niger* (among others), has been identified as the most efficient fungi for LEVO assimilation in a wide range of pH and temperature (Xie et al., 2006). *Aspergillus niger* has been detected in DE soil from Amazon (Glaser, 2007; Grossman et al., 2010; Lucheta et al., 2016), but the lack of studies focused on the characterization of microorganisms in DE from Europe do not allow to accept or exclude this hypothesis.

The presence of degradation phenomena acted by bacteria was also hypnotized in the previous section 4.1.1. In fact, the distributions of even and odd *n*-alkanes could be affected by biodegradation (Chandra et al., 2013; Daccò et al., 2020). Moreover, the prevalence of even carbon number linear alkanes could be influenced by the presence of petroleum derivates such as bitumen or tar (Whelton et al., 2021).

The absence of meaningful correlation between PAHs, anhydrous sugars, charcoal fragments, and Raman parameters could indicate that different biomass burning products undergo different degradation and/or transport mechanisms in the soil.

5 Fecal sterols and stanols

In this chapter the results of the analysis of fecal biomarkers are reported. In the first part of the chapter the focus is on sterols and stanols referred to the fresh fecal samples of animals. These were, in fact, used to build a database of fecal proxies related to different sources of fecal pollution. Beginning from these concentrations, the data exploration started following the approaches most frequently used and reported in literature, then different statistical models were tested to investigate the possible identification of the fecal input in samples for which it is possible to hypothesize a mixed origin.

The second part of the chapter is focused on the results of the fecal impact attribution on the samples from Dark Earth (DE) layers.

5.1 Sterols in animal feces

As reported in section 2.6, fresh fecal samples from different species were sampled. Table 5.1 reports a summary of the characteristics of animal feces and the number of samples collected for each animal.

Table 5.1. Samples collected for the determination of fecal sterols and stanols divided by species and group.

Animal	Group (i.e., genera/subfamily)	# of samples
Donkey	Equine	17
Horse	Equine	12
Sheep	Caprine	15
Goat	Caprine	12
Cow	Bovine	15
Pig	Swine	12

Firstly, the sterols and stanols fingerprints for the six classes of animals were investigated. The global concentrations of the sterols and stanols spanned between 0.35 µg g^{-1} and 3.2 mg g^{-1}.

The results are reported in Table 5.2 as mean concentration ± 95% confidence interval of each analyte and each animal, and in Figure 5.1 as bar plots of the mean values of the analyte concentration normalized for the total sterol content, where error bar represents the 95% confidence interval of each distribution. The abbreviations of the sterols and stanols are reported in the previous Table 1.3.

Similarities in the distribution of analytes were found for animals belonging to the same family, such as donkey and horse (Equine genus), and sheep and goat (Caprine subfamily). The most abundant congener for all the herbivores (pig group is excluded) is 24-ethyl-coprostanol (24COP), whereas for pig the most abundant analyte is coprostanol (COP). For donkey, horse, sheep, and goat 24COP represents more than 30% of the total congeners (Figure 5.1).

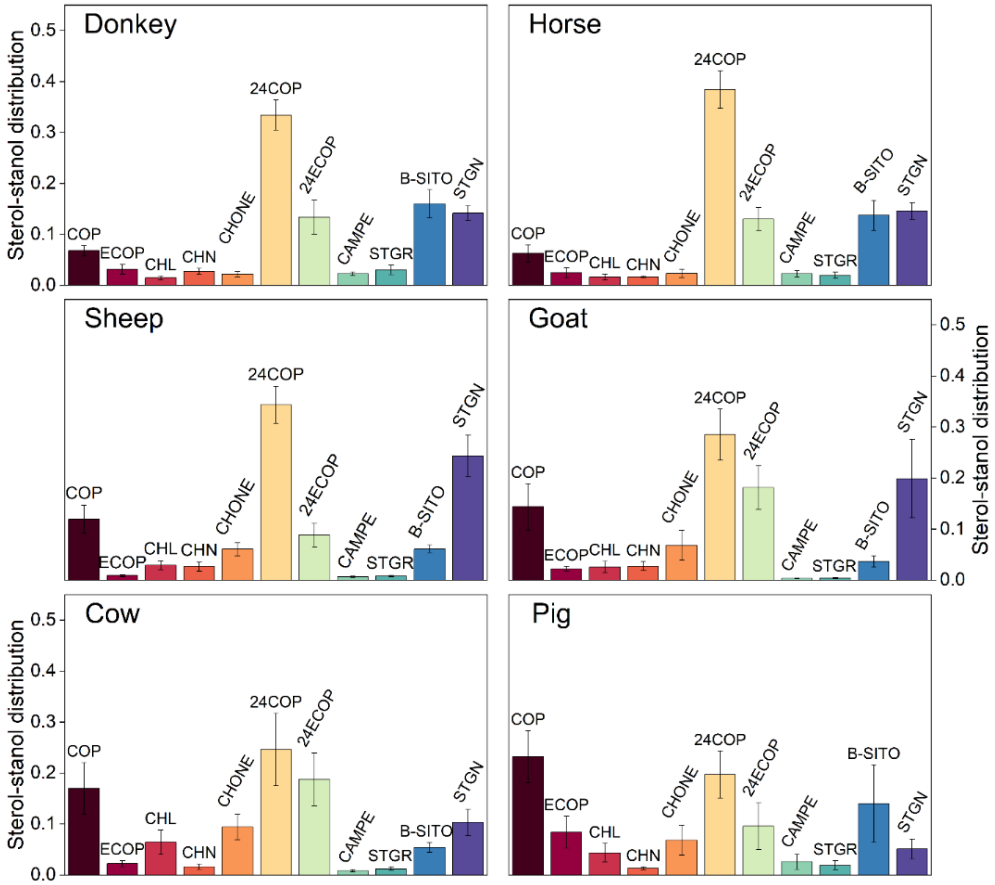

Figure 5.1. Sterols and stanols distribution of the feces of the six animal groups investigated. Concentrations were normalized on the total sterols and stanols content. Columns represent the mean values for each analyte, error bars are referred to the 95% confidence interval of the distribution.

The distributions of donkey and horse are dominated by 24COP (with concentrations of 700 ± 300 μg g^{-1} and 1100 ± 300 μg g^{-1}, respectively), while other abundant congeners are the C_{29} sterol β-sitosterol (β-SITO) (respectively 300 ± 90 μg g^{-1} and 340 ± 90 μg g^{-1}) and stanol stigmastanol (STGN) (respectively 300 ± 100 μg g^{-1} and 400 ± 120 μg g^{-1}). Sheep and goat present 24COP concentrations of 1300 ± 500 μg g^{-1} and 1400 ± 600 μg g^{-1}, respectively; the second most abundant analyte is STGN (with concentrations respectively of 850 ± 280 μg g^{-1} and 910 ± 500 μg g^{-1}), whereas the third major component is COP for sheep (400 ± 130 μg g^{-1}), and 24-ethyl-epi-coprostanol (24ECOP) for goat (900 ± 500 μg g^{-1}).

Sheep and goat samples are defined by an elevated amount of stigmastanol, but lower β-sitosterol, if compared to *Equine*. For *Caprine*, coprostanol represented more than 10% of the total sterols, presumably reflecting the less selective diet of sheep and goats compared to the one of donkeys and horses.

The fecal fingerprint of pig is the most homogeneous as it presents four species between sterols and stanols representing ~70% of the total sterols/stanols content. Indeed, COP, 24COP, β-SITO, and 24ECOP are present in the concentration of 320 ± 70 μg g^{-1}, 300 ± 150 μg g^{-1}, 80 ± 20 μg g^{-1}, 160 ± 110 μg g^{-1}, with coprostanol showing the highest concentration because of the omnivorous diet. The results referred to pig are aligned with the ones reported by Harrault et al., (2019).

The sterols and stanols distribution of cow does not show any predominant (>30%) compound. As previously observed, the dominant stanol is 24COP (350 ± 200 μg g^{-1}) followed by 24ECOP (200 ± 80 μg g^{-1}). The third most abundant stanol is COP (180 ± 80 μg g^{-1}). The raise in bovine content of coprostanol with respect to the other herbivores could be due to the feed improvement with soy protein; indeed it has been reported that a diet enriched in soy brings to a higher conversion of cholesterol in coprostanol (Tanaka et al., 1983). In fact, even if the fecal samples were collected in family conduction dairies and natural lodges, it has been told by farmers that the diet is often improved with soy-derived food preparation.

These abundances are concordant with the literature, where the feces of herbivores are reported to be richer in stanols and sterols with 29 carbon atoms (typical sterols and stanols of vegetation), whereas omnivore feces are richer in C_{27} congeners, representative of a diet that includes also complex proteins and elevated amounts of cereals (Derrien et al., 2017, 2011; Evershed et al., 1997; Jardé et al., 2007a; Leeming et al., 1996; Shah et al., 2007; Tyagi et al., 2008). The presence of abundant quantities of 24ECOP in the fecal fingerprint of all the studied animals is in contrast with the

interpretation of the source epimer proposed by Birk et al., (2011), Bull et al., (2002) to be solely from microbial digestion in soil. A similar pattern, with elevated content of 24ECOP, was determined also by Harrault et al., (2019). The presence of relevant amounts of 24ECOP could be indicative of processes of epimerization occurring in the gut of mammals too.

Table 5.2. Concentration of sterols and stanols for the six animal feces analyzed the values represent the mean values and 95% confidence interval.

animal	COP	ECOP	CHL	CHN	CHONE	24COP	24ECOP	CAMPE	STGR	β-SITO	STGN
Donkey	130 ± 50	60 ± 30	25 ± 8	50 ± 20	40 ± 20	700 ± 300	280 ± 140	40 ± 10	50 ± 20	300 ± 90	300 ± 100
Horse	170 ± 80	60 ± 30	45 ± 20	40 ± 10	70 ± 30	1100 ± 300	330 ± 100	55 ± 15	45 ± 10	340 ± 90	400 ± 120
Sheep	400 ± 130	30 ± 10	90 ± 30	100 ± 50	200 ± 70	1300 ± 500	340 ± 140	23 ± 9	30 ± 10	230 ± 80	850 ± 280
Goat	580 ± 230	100 ± 40	90 ± 40	130 ± 60	250 ± 100	1400 ± 600	900 ± 500	13 ± 5	14 ± 5	160 ± 70	910 ± 500
Cow	180 ± 80	20 ± 7	50 ± 15	20 ± 10	90 ± 30	350 ± 200	200 ± 80	9 ± 5	10 ± 3	55 ± 20	130 ± 100
Pig	320 ± 170	130 ± 70	30 ± 10	16 ± 8	90 ± 50	300 ± 150	160 ± 110	15 ± 5	12 ± 3	80 ± 20	60 ± 40

5.2 Sterols and stanols diagnostic ratios

Many ratios between sterols and stanols have been proposed as proxies to determine fecal contamination sources (Derrien et al., 2017). Some of these diagnostic indices have been calculated starting from the sterols and stanols concentration determined in the feces of the animals (Table 5.3). The dataset was improved by adding observations of human feces from Leeming et al., (1996) and Shah et al., (2007). Unfortunately, it was impossible to find references reporting the concentration of 24-ethyl-epi-coprostanol for human feces. Hence the ratios that use this stanol are excluded from further discussions. The full dataset is available in Appendix E. Figure 5.2 represents the boxplots of the ratios calculated for each animal group.

Table 5.3. Diagnostic ratios of sterols and stanols with threshold and resources.

Ratio #	Calculation	Diagnostic threshold	Source	Reference
R1	$\frac{24COP}{\sum(total\ sterols)}$	> 0.5 – 0.6	Herbivore	Isobe et al., 2002
R2	$\frac{CAMPE + \beta - SITO}{CHL}$	> 2.5 > 1.5	Bovine Swine	Jardé et al., 2007a, 2007b
R3	$\frac{COP + ECOP}{CHL}$	> 3.7 < 0.7	Swine Bovine	Jardé et al., 2007a, 2007b
R4	$\frac{\beta - SITO}{COP}$	2.3 < R6 < 3.3 0.2 < R6 < 0.7	Bovine Swine	Gourmelon et al., 2010
R5	$\frac{COP}{CHL}$	> 1	Human	Quéméneur and Marty, 1994
R6	$\frac{COP + ECOP}{COP + ECOP + CHN}$	> 0.7	Human	Bull et al., 2002

For R1 all the distributions are below the threshold set at 0.5. The region above this line should be representative of the herbivore fecal pollution, but none of the samples analyzed (all below 0.5) would be associated to herbivorous fecal input using this ratio.

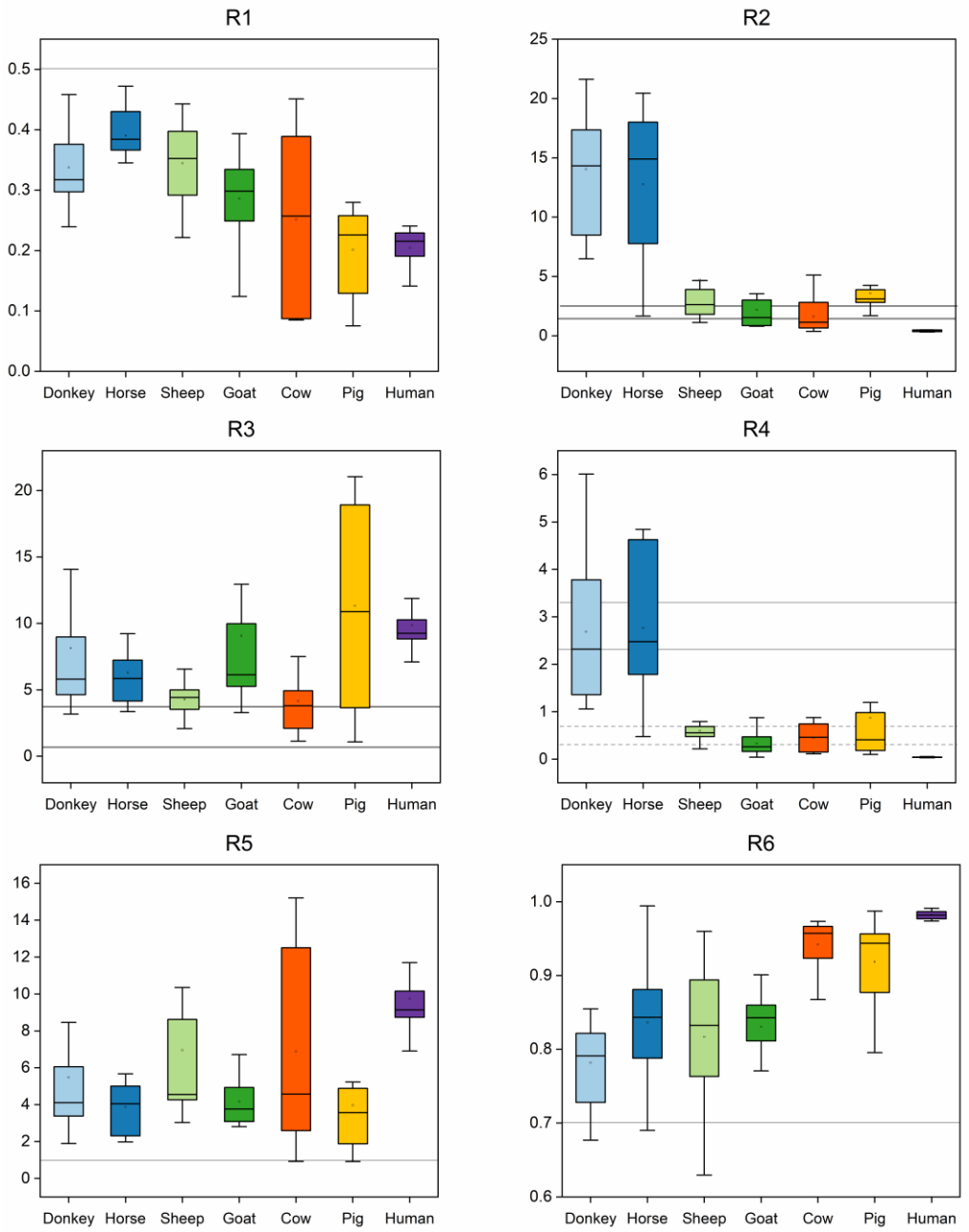

Figure 5.2. Boxplot of the diagnostic ratios for human and animals' sources. Box limits represent the 25%-75% of the distribution, whiskers define the range within the 1.5 interquartile region of the distributions. Black squares represent the mean values, lines the median values. Lines indicate the thresholds.

R2 was proposed to identify different animal input from bovine and swine, according to the content of plant and animal sterols. The thresholds are R4 > 2.5 for bovine pollution and R4 > 1.5 for pig. As visible from Figure 5.2 the distributions of sheep, goats, cows, and pigs falls approximately in the same region: between 1 and 4; additionally, the distribution related to pigs falls above 2.5, indicated to be representative of bovine feces, whereas cow distribution is closer to the threshold

linked to pig presence (1.5). The distribution related to the humans falls below the threshold identified for pigs, indicating that this ratio would be possibly used to distinguish human feces, enriched in CHL, from the animal ones.

R4 is built similarly to R2 (plant sterol/animal stanol) and was formulated to distinguish the fecal input of pigs and cows. For R4 the pig distribution is located precisely in the interval suggested by Gourmelon et al., (2010) for pig identification (solid lines in R4 graph in Figure 5.2). However, the distributions of cow, goat and sheep fall in the same interval. R2 and R4 results are essentially indicative of the same phenomenon: horses and donkeys have higher values, suggesting that feces are mostly affected by plant sterols and stanols, but all the other distributions cannot be distinguished using these ratios. Also in this case, the human distribution falls below all the other ones, because of the higher content of cholesterol-derived compounds, namely coprostanol. Additionally, the R4 box-plot, displays the interval lines proposed by Gourmelon et al., (2010) to distinguish cow feces. Even if the cow distribution does not fall into the proposed area, mean values of the distribution of equines are around the proposed values ($2.3 < R4 < 3.3$), suggesting the suitability of this ratio for the identification of those animals.

R3 was proposed by Jardé et al., (2007a, 2007b) to differentiate between pigs and cows input in soil. Values above 3.7 and below 0.7 should represent pig and cow feces, respectively. Figure 5.2 highlights the higher mean value obtained for pigs and humans with respect to all the other animals. However, the distributions referred to the other animals partially overlap with the values calculated for pigs, possibly due to the large variability of pig samples. Nevertheless, mean values for all the animals are above the threshold characteristic of pigs.

R5 and R6 were formulated to distinguish human from animal fecal pollution. The thresholds 1 and 0.7 are represented with a solid line in Figure 5.2 for R5 and R6, respectively. These ratios for all the animals have values above the identified thresholds (except for a few observations represented with the interquartile whiskers). This means that the feces from all the animals would be interpreted as human feces when those specific values are used. In turn, mean values of humans fall above the limits proposed by Quéméneur and Marty, (1994) and Bull et al., (2002). Hence, differentiations between the classes can be evidenced when using these ratios but with a careful interpretation of the thresholds.

As evidenced in this section the use of the ratios from literature is misleading for the discrimination of fresh feces of different animals. During the last 20 years some researchers highlighted the scarce classification strength of these ratios, and studies have been conducted to find more reliable methods for the distinction between fecal inputs.

Mudge and Duce (2005) studied the lipid content in the aquatic environment to identify the most polluted sites using the principal component analysis (PCA) and the partial least squares analysis (PLS) for the first time. They started from the analysis of raw data (concentration of the single analytes), that were converted into proportion on the total sterols/stanols to stabilize the concentration fluctuations. From the PCA results, they were able to locate the principal input of pollution, and the sterols and stanols were sorted into sources from prior knowledge. As a result, PLS was used to classify sewage samples starting from the chemical signature of specific references.

A different application of the pattern recognition techniques was the one reported by Saim et al., (2009). They determined the concentration of sterols and stanols in wastewater samples from different sources and farms. Their interest was focused on cluster analysis (CA), discriminant analysis (DA), and PCA to investigate the most powerful variables for animal clustering and the identification of fecal pollution sources. The cluster analysis of the variables identified three groups: (i) coprostanol; (ii) cholesterol and β-sitosterol; (iii) stigmasterol and stigmastanol. DA allowed the clear distinction of four animal groups (humans, chickens, quails, and horses) and suggested that the variables with the highest discriminative power were cholesterol, coprostanol, and β-sitosterol. In the work by Saim et al. (2009), the explained variance of the PCA was higher than that of Mudge and Duce (2005), and the variables identified to be the most representative for the separation into groups were the same ones revealed by the DA.

Studies by Martins et al., (2011, 2007) reported the results of the sterols and stanols analysis in lagoon and estuarine environments for a deeper comprehension of the mechanisms of fecal, algal and terrestrial pollution in water. The initial use of sterols and stanols ratios for the identification of major pollution inputs in water was described. Specifically, the ratio [(COP + ECOP)/total sterols] allowed to determine the sites where the fecal pollution was prominent (Venkatesan and Kaplan, 1990), while the ratio ECOP/COP was used to establish the degree of sewage treatment: ratios higher than 0.8 indicated an abundant conversion degree and meant that the receiving water was treated sewage (Mudge and Seguel, 1999). Secondly, the PC1 and PC2 extracted

from the PCA highlighted two different phenomena affecting the studied sites: PC1 differentiated highly contaminated sites from less contaminated ones, while PC2 discriminated the fecal input from the plant/terrestrial input. This study correlated fecal contamination to the presence of coprostanol and epicoprostanol, excluding the possibility for phytosterols (such as β-SITO, STGN, STGR, and CAMPE) of coming from animal feces.

The use of ratios for pollution sources discrimination has been questioned by Derrien et al., (2011), studying cow and pig fresh feces, manure, and treated and untreated slurries. The study focused on the distribution of sterols and stanols in different samples: feces from bovines were mostly described by C_{29} steroid congeners (24ECOP and β-SITO as more abundant), whereas pig excrements were defined by a higher contribution of 3β and 5β stanols (COP and 24COP). Besides, the authors reviewed the main ratios used in literature to discriminate feces pollution sources, highlighting the lack of specificity of almost all of them as due to the different metabolism of the animals in different phases of their life more than to difference in diets and/or sewage treatments. Finally, the study exploited the PCA analysis as an alternative method to fix the lack of discrimination evidenced by sterols and stanols ratios. The results showed that PCA is a valuable tool even when the samples were affected by different ageing/treatments. This approach was further developed using samples of human feces and subjecting the samples to a dilution/runoff effect (Derrien et al., 2012). Even in this case, the PCA resulted in a suitable technique to determine the multivariate clustering of the samples. Other successful applications of PCA to sterols and stanols analysis regarded the presence of classes with both spatial and seasonal differentiation (de Abreu-Mota et al., 2014; Derrien et al., 2015).

The combined use of PCA and CA was proposed by Harrault et al., (2019) as an optimized classification method to differentiate fresh faces of both omnivores (pigs, dogs, humans) and herbivores (cervids, ovine, horses, cows, lemmings). The resulting model was used on archaeological samples to infer the input sources of past settlements.

Linear regression models were applied by Tyagi et al., (2007) to farm animals. In this study, the concentrations of fecal biomarkers were analyzed and correlation coefficients between variables were determined. The dataset was split into two: the training set and the test set, the predictive efficiency was around 98-100%, and some observations from run-off samples were examined to evaluate the performances of the model, resulting in 100% of the samples being efficiently predicted. The research did

not investigate the source identification strength of the model when mixed fecal inputs were present in the environment.

A different strategy, involving the discriminant analysis, was proposed by Shah et al., (2007) for the classification of animals by their feces sterols content. The concentration of sterols and stanols in feces from different herbivores, birds, carnivores, and humans was determined. Although the sterol fingerprint permitted a good separation into the four groups, the interest of the authors was also in the identification of different contributions when the sources are mixed. Even though the initial DA model run on reference mixtures failed the prediction, an optimized model was built to distinguish only two groups: one including human contribution, and a second one containing the other combinations of animals (excluding humans). In this case, the model was able to classify the samples where the human contribution was present, but no other information could be achieved (i.e., the other animal present in the mixed observation).

In conclusion, the literature shows that specific sterols and stanol ratios are helpful when the question is how different animal distributions behave. Additionally, the multivariate statistical analysis approach is the most reliable to identify the most significant variables for fecal source identification. Whereas pattern recognition analysis is not strictly informative of the different inputs in mixed samples. It is evident that, at the present time, there is no adequate method for interpreting data displaying such a high variability.

5.3 Multivariate analysis

Using the discussed literature as a guide for the data exploration, multivariate analysis was performed on the raw data from animal feces.

Firstly, hierarchical cluster analysis was carried out to determine the dependence between the variables. Figure 5.3 represents the dendrogram of the variables. Ward metric was chosen to calculate the similitude matrix. Figure 5.3 shows three groups cutting the dendrogram at 1.5. The most distant group is formed by CHL and cholesterol-derived analytes (COP, CHL, CHONE, and ECOP). The second group of three variables is formed by CHN, STGN and 24COP that, as already reported in section 1.2.5 and Figure 1.6, derive from sterol precursors by reduction reactions performed by gut microbiota (Leeming et al., 2014). The third group is formed by three

C_{29}-sterols: CAMPE, β-SITO and STGR, representing the plant sources and reported to be characteristic of diets with a small rate of conversion from Δ5-sterols to 5β-stanols (Leeming et al., 1996).

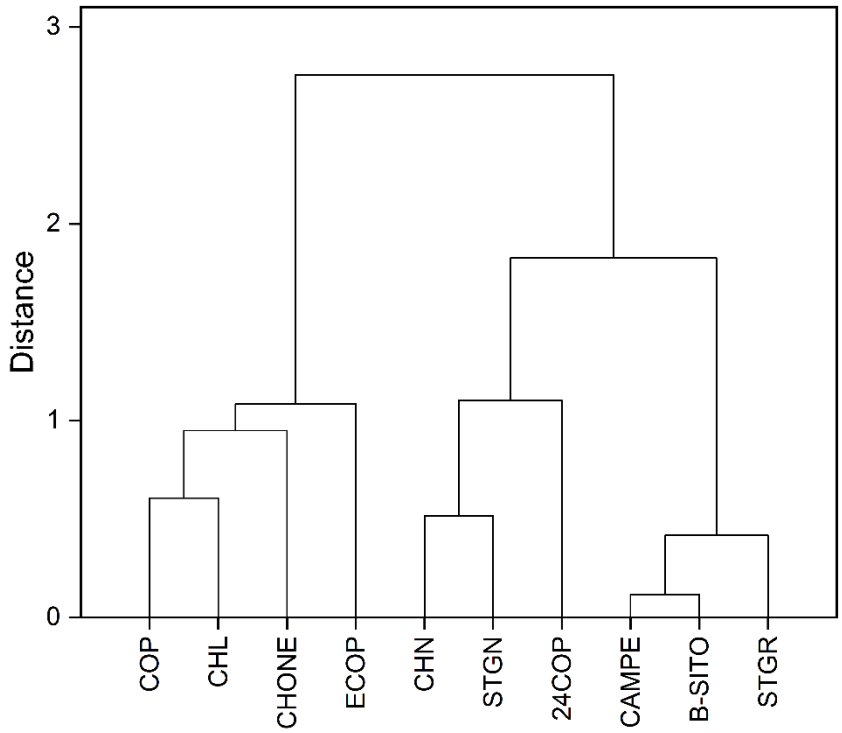

Figure 5.3. Dendrogram resulted from the CA of variables.

PCA was also performed on the raw data acquired on animal feces using all the variables. The cumulative variance of the first three PCs is 69.7%, each one is responsible, for the 30.4%, 25.3%, and 14.0% of the variance, respectively. From the extracted eigenvalues it is possible to identify the variables that account positively or negatively on the PCs from their coefficients (Table 5.4).

Table 5.4. Coefficients associated to all the variables for the first three PCs.

Variable	PC1	PC2	PC3
COP	-0.35	0.35	-0.12
ECOP	-0.17	0.30	-0.10
CHL	-0.15	0.32	0.48
CHN	-0.02	-0.29	0.47
CHONE	-0.39	0.23	-0.10
24COP	0.14	-0.35	-0.57
24ECOP	-0.18	-0.14	0.29
CAMPE	0.44	0.29	0.13
STGR	0.44	0.24	0.06
β-SITO	0.48	0.23	0.04
STGN	0.03	-0.46	0.30

The variables with a higher weight on the first PC are the vegetal sterols CAMPE, STGR and β-SITO, namely the same ones that constitute the most distant group identified by CA (Figure 5.3). In parallel, CHONE and COP are the variables with the smallest coefficients, as these represent, also in the dendrogram, the counterweight to the plant sterols.

The coefficients of the variables specify the direction of information maximized by the PC. Indeed, PC1 can discriminate the observations by the content of plant sterols (with greater scores of PC1) and of cholesterol-derived compounds (showing lower PC1 scores). Hence, PC1 classifies the observations by the content of non-processed plant sterols.

PC2 is influenced positively by COP and CHL, and negatively by STGN and 24COP. The latter two are 5α-stanols typical of the conversion of plant sterols, although COP is a 5β-stanol derived from the transformation of cholesterol. Thus, PC2 can inform about the degree and type of gut conversion: herbivores will have lower PC2 scores whereas omnivores are expected to have higher PC2 scores.

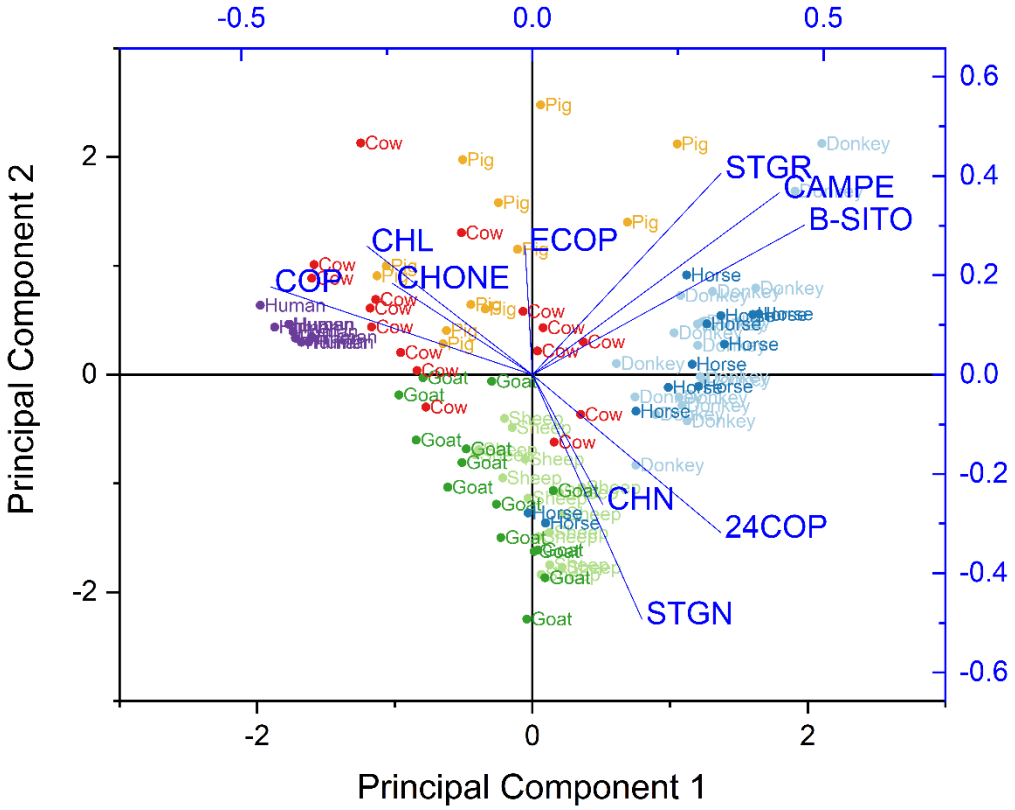

Figure 5.4. The biplot of the PCA performed on the fecal sterols and stanols determined in the samples of feces from animals.

Figure 5.4 is the biplot of the PCA performed and reports PC2 vs PC1. The loading plot informs about three groups of variables: (i) the plant sterols group (CAMPE, STGR, and β-SITO); (ii) the cholesterol-derived compounds (CHL, ECOP, COP, and CHONE); (iii) and lastly the group constituted by 24ECOP, CHN, STGN, and 24COP. The observations are arranged into three main groups partially overlapped: donkey and horses are located mostly in the right-hand half of the plot, meaning a high presence of sterols of vegetal origin; goats and sheep are grouped in the bottom left quarter, with a lower vegetal sterols contribution and a higher input of processed vegetal stanols. Observations regarding cows and pigs are located in the top-left region, close to humans. This area is determined by greater contents of the cholesterol-derived compounds (with few pig observations on the top right quarter and cow observations in the bottom quarters).

5.4 Classification models

The first published scientific method of classification is the one proposed by Carolus Linnaeus to classify all the living species; since then, statisticians and mathematicians implemented and improved classification models. The sophistication of the classification methods rose exponentially since the first mathematical methods developed at the beginning of the 20th century by Sir R.A. Fisher and colleagues (McLachlan, 2005), up to the most recent machine learning methods.

Classification models are usually based on the reiteration of a function or algorithm able to differentiate objects. Generally, the classification models are built on a portion of the interesting dataset, called training set. Consequently, the model is validated on the leftover observations of the dataset, this is called test set.

The classification models that will be studied in the next paragraphs give the probability associated to each class as output of the prediction. In other words, once the model is built on the training set, the r function `prediction()` applied on the model parameters achieved and on the test set gives, as output, a probability associated to each class of the model for the observation tested.

For the fecal proxies the goal was to find a suitable model to predict the type of fecal input, to be applied on the DE samples. To achieve this aim, the focus was primarily

put on the best classification method for animal feces observations. Before performing the models, the sterols and stanols distributions were normalized on the total amount of congeners determined in each sample. The dataset is available in Appendix E.
The different models were built under the same conditions and the performances of the prediction were investigated changing several parameters such as:

a) The classification levels: classification of the animals (donkeys, horses, sheep, goats, pigs, cows, and humans) was performed, as well as grouping for genus/subgroup (equine, caprine, swine, bovine, and homo).
b) The number of cycles of the model: under the same conditions (classification parameter and training set largeness) the number of times that the model recalculates itself was increased to estimate how the accuracy of the prediction match changes.
c) The number of observations used to build the model: the largeness of the training set was varied to evaluate the robustness of the model and the classification match performances of the prediction.

To simulate the possible input of different fecal sources together, a dataset based on different binary or ternary mixtures at variable percentages (for example, 20% of donkey feces + 80% of sheep faces) was created. All the possible combinations were calculated, and the predictive power of the model was tested on this new set of data.

5.4.1 Linear Discriminant Analysis (LDA)

LDA and PCA are defined as models for dimensionality reduction. The LDA technique was developed to transform the data elements into a lower dimensional space, which maximizes the class separability (Tharwat et al., 2017). LDA basically identifies the hyperplane in which the differentiation between data classes is maximized, minimizing the spread within each class. This building mechanism makes LDA suitable for solving classification problems (Ioffe, 2006).
The idea behind the method was to build a subspace of fewer dimensions (compared to the original data sample dimension) in which the data points of the original problem are separated. Separability is defined in terms of statistical measures of mean value and variance (Xanthopoulos et al., 2012).
To pursue this goal, the model operates in three steps: (i) definition of the between-class matrix, based on the separability (i.e., the distance between the means) of the

classes; (ii) definition of the within-class matrix, in which the difference between all the samples and the mean of the group, is calculated; (iii) construction of the space in which the between-class variance is maximized and the within-class variance is minimized. As schematized in Figure 5.5, the separability of two groups A and B is maximized, and the variability within the subgroup into a lower dimension space is minimized.

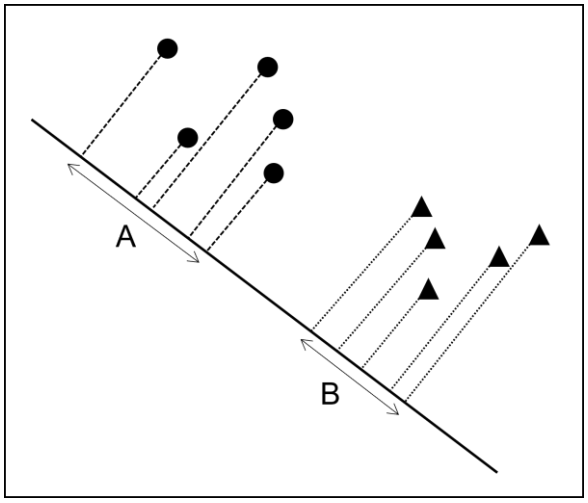

Figure 5.5. The basic concept of LDA: the observations are represented in two dimensions and associated with two groups A and B, the one dimension (line) on which projecting the objects is the representation of the basic concept of the optimized lower dimensions subspace of the LDA (elaboration from Xanthopoulos et al., 2012).

The application of LDA to the classification problems is associated to the possibility to have a separability parameter that can be used to classify the test observations. Indeed, LDA allows finding the population that maximizes the likelihood of the test observations, and in the case of a prior knowledge of the proportion of each group in the present, the model is able to maximize the posterior probability that a test observation belongs to a particular population group (Guo et al., 2007).

5.4.1.1. LDA validation

LDA was applied to the dataset of fecal sterols and stanols content of animals and humans, the CRAN package used was `MASS`. Figure 5.6 represents the performances of the LDA model when the training set consists of 85 observations and the test set is a random choice of 11 objects. The repetition of the same model with a different random choice of the 11 objects test set consisted of an increasing number of cycles: 1 cycle, 10, 100, and 1000 cycles. For each number of cycles, the output was a percentage

value of correct prediction. The percentage mean and the variability of prediction match were calculated on the repetition of the same number of cycles for 50 reiterations. Table 5.5 reports the percentage of match and the variability (expressed as the standard deviation) as the classification parameters change.

Table 5.5. The percentage of match of LDA classification model changing the classification parameter and the number of cycles.

Classification	1 cycle	10 cycles	100 cycles	1000 cycles
Animal	64 ± 14%	70 ± 4%	68 ± 1%	67.6 ± 0.5%
Group	88 ± 8%	87 ± 2%	87 ± 1%	87.0 ± 0.3%

Table 5.5 shows the stabilization of the percentage of match at 100 cycles: the mean, for animal level of classification, is 68 ± 1% at 100 cycles and 67.6 ± 0.5% at 1000 cycles. Conversely, the classification based on groups is more stable, showing similar percentages of match at 10 cycles already (87 ± 2%); by increasing the number of cycles, the effect of cycle reiterations affects the variability, that decreases.

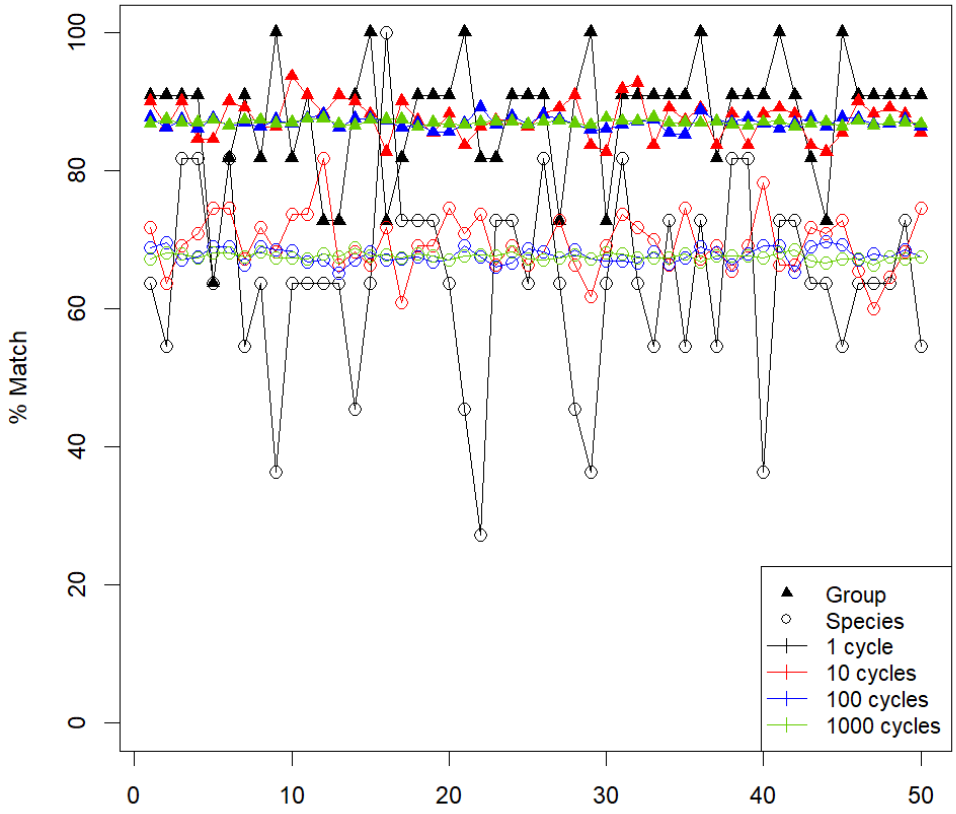

Figure 5.6. Match percentage of LDA varying the number of cycles for the validation and the classification parameters. Dots represents the classification for animals (donkeys, horses, sheep, goats, pigs, cows, and humans) triangles stand for superior group classification (equine, caprine, swine, bovines, and humans). Colors represent the change in the number of cycles.

As shown in Figure 5.6 and Table 5.5, the classification parameter affects the performances of the prediction match in a dramatic way: the group classification shows percentages of match always above 85%, whereas the percentage of match using the animal classification is between 65 and 70%. Moreover, a decreasing variability is observed with the increase of the number of cycles, as displayed in Figure 5.6, as the blue and green distributions (referred to 100 and 1000 cycles) are much more constant than the black and red ones (1 cycle and 10 cycles). For example, the distribution with the highest variability (classification of animals with 1 cycle, 64 ± 14% of match) presents extremes of ~27% and 100% (Figure 5.6). This means that in the first case (~27%) only 3 out of 11 observations of the test set were correctly predicted; in the second case, objects of the entire test set were perfectly classified.

The performances of LDA were investigated as the numerosity of the test set increases while the training set, used to build the model, decreases. Hence, the percentage of match will reflect the prediction capability of the model in the limit conditions: starting from 1 object in the test set (i.e., 95 objects in the training set), to 81 observations in the test set (i.e., 15 objects in the training set), with a step of 10. For this second test, the number of cycles for the validation was set to 100, considered as a suitable value for the correct evaluation of the match % (see above), and the number of reiterations at 50.

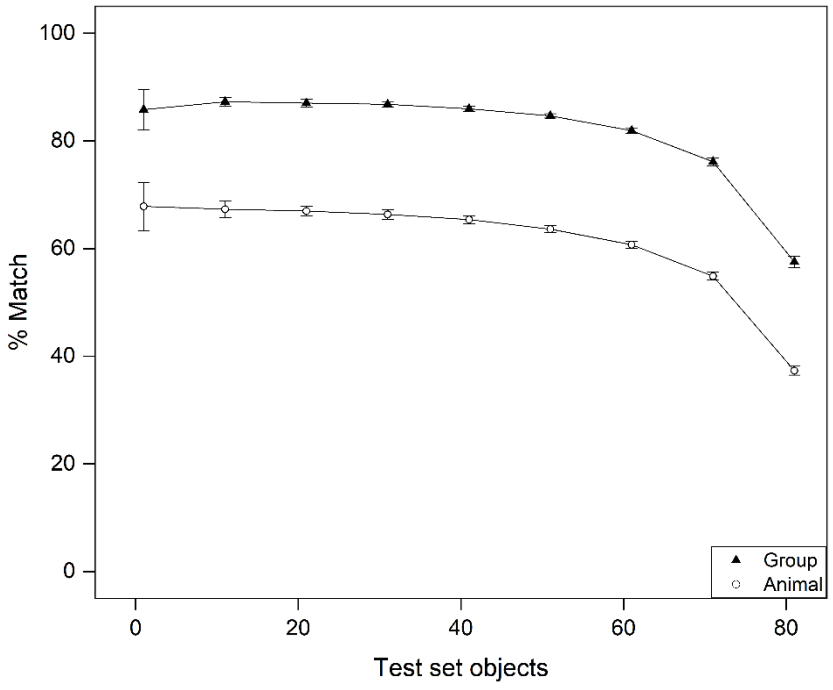

Figure 5.7. Percentage of match of the prediction of LDA model varying the number of objects of the test set. Triangles and dots represent the group and the animal classification, respectively.

Table 5.6 and Figure 5.7 report the mean value and the standard deviation for each distribution of the match calculated on the distribution of 50 validation cycles reiterations.

The similar shape of curves in Figure 5.7 indicates that the threshold of the training set numerosity for the match percentage stability is the same for both the classification parameters. Indeed, the match is above 60% for animal classification, and above 80% for group classification, until 61 (i.e., training set has 24 objects). Above this value, the % match of both curves decreases at higher rate (steep slope decrease), reaching the last point at 37.3 ± 0.8% for animal classification and 57 ± 1% for group classification. For both the classification levels, the net difference in the match prediction from the first to the last point is around 30% (30.7% for animal classification, 29% for group classification).

Except for the first point, the variability is constant through the decreasing of the training set abundance, and its value is the same as the previous test with the same number of cycles reported in Table 5.5.

Table 5.6. Percentage of match changing the number of objects of the test set. Values refer to the mean value of 100 cycles of the same model and 50 validation reiterations ± the standard deviation of the distribution.

Test set objects	Animal classification	Group classification
1	68 ± 4%	86 ± 4%
11	67 ± 1%	87.2 ± 0.8%
21	67.0 ± 0.9%	87.0 ± 0.7%
31	66.3 ± 0.9%	86.8 ± 0.5%
41	65.4 ± 0.7%	85.9 ± 0.5%
51	63.6 ± 0.6%	84.6± 0.4%
61	60.7 ± 0.7%	81.9 ± 0.5%
71	54.9 ± 0.7%	76.1 ± 0.7%
81	37.3 ± 0.8%	57 ± 1%

5.4.2 Trees

Trees are data mining models based on the classification of observations by specific item values that can be used to classify and predict. The implementation of the model is based on the recursive application of an algorithm for partitioning a recorded dataset (Jadhav and Channe, 2016). Different algorithms permit to build trees based on the linear regression when the variables are numeric (regression trees) or more general classification trees. Tree models were introduced in the 1960s and underwent rapid diffusion in several disciplines because of the ease of use and setting, coupled with a relative robustness and a clear graphical representation (Song and Lu, 2015).

The decision trees are flow chart-like structures with roots, nodes, and leaves (terminal nodes). Each node indicates a test condition of a specific attribute: root nodes represent a choice that introduces to the subdivision into two or more subsets, each internal node represents a choice that can be performed at that point of the structure, while terminal nodes represent the final result of a combination of decisions or events and are named following the classification labels. Branches represent the result of the upper test condition of the node, and each path from the top to the bottom of the decision tree constitutes a decision rule. These decision tree pathways can be traduced into an 'if-then' rule. For example: *"if condition 1 and condition 2, (...), and condition n occur, then outcome j occurs"* (Hastie et al., 2021; Jadhav and Channe, 2016; Song and Lu, 2015).

The typical representation of a classification problem solved with the decision tree is reported in Figure 5.8 (Loh, 2014) together with the scatter plot of the observations where lines represent the variables thresholds applied to distinguish among species of iris flowers (*Versicolour*, *Virginica*, and *Setosa*). In the decision tree, the numbers of misinterpreted data are reported under the leaf nodes. In fact, some *Versicolour* and *Virginica* observations overlap (as scatter plot).

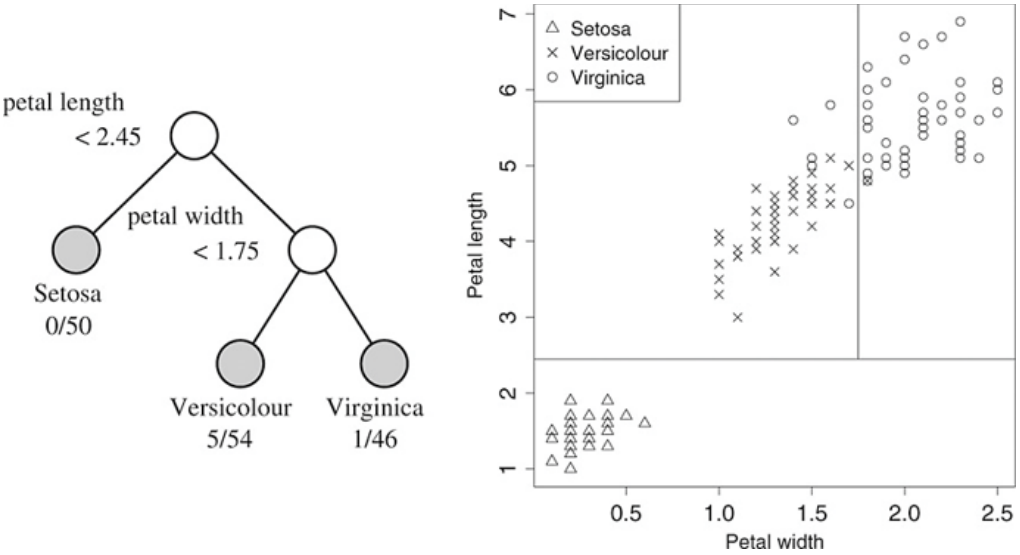

Figure 5.8. Classification of three species of iris flower following the petal width and length. Both the decision tree and the scatter plot of the observations are reported (Loh, 2014).

5.4.2.1 Tree validation

The classification tree model was applied to the animal/human feces dataset using the CRAN package `tree`, and the representations of the trees were made with the CRAN

package `rpart.plot`. Figure 5.9 reports the examples of two classification trees with the two levels of classification: animals and groups (genus/subfamily). As evident, only the most significant eigenvalues reported in the PCA (section 5.3) were selected as variables for the classification: β-SITO, CHN, COP, ECOP, STGR, and STGN.

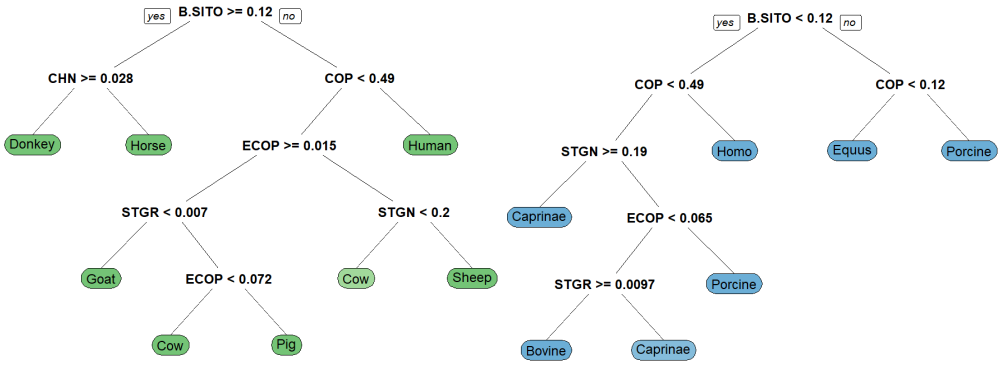

Figure 5.9. Classification tree of the observation classified for both the levels: animals (green nodes) and groups (blue nodes).

The same approach followed for LDA models was applied to the classification tree. Figure 5.10 represents the prediction match of tree when the training set consisted of 85 objects and the test set was a random choice of 11 elements. The number of cycles of the model was increased from 1 to 1000 (steps of one order of magnitude). Also in this case, the mean values and standard deviation were calculated on the percentage of matching derived from 50 reiterations of the same validation cycle (i.e., 50 reiterations of the 1 cycle model, 50 reiterations of the 10 cycles model, etc.).

The mean values and variability (expressed as the standard deviation) of the percentage of prediction match, as the classification parameters change, are reported in Table 5.7.

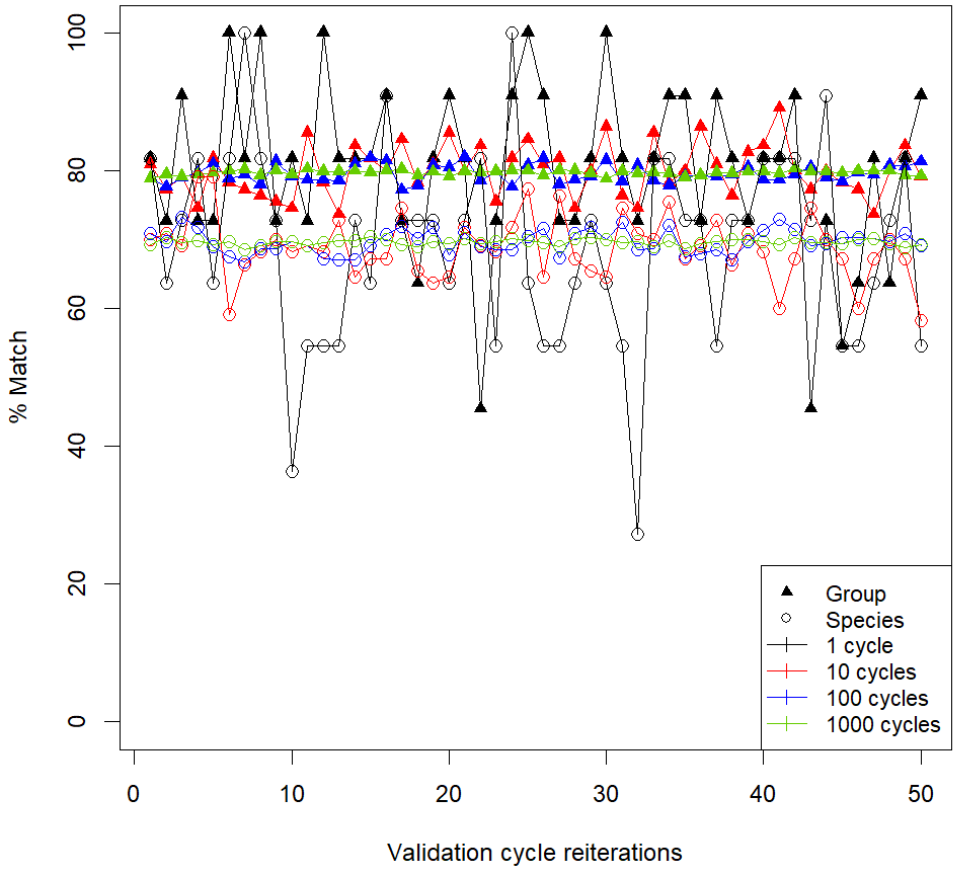

Figure 5.10. The percentage of match of the prediction using tree model varying the number of validation cycles (1, 10, 100, 1000). Empty dots are referred to the classification level "animal", whereas triangles indicate the classification level "group".

Table 5.7. Mean and standard deviation of prediction match of classification tree model changing the classification parameter and the number of cycles.

Classification	1 cycle	10 cycles	100 cycles	1000 cycles
Animal	69 ± 14%	69 ± 5%	70 ± 2%	69.9 ± 0.4%
Group	80 ± 12%	80 ± 4%	80 ± 1%	79.7 ± 0.3%

When the classification level is 'animal' the match of the prediction is around 70%, whereas when the classification is made for the groups the match percentage increases up to 80%.

The prediction match percentage of the largest distribution (namely the 1 cycle model based on animal classification) is 69 ± 14% (Table 5.7), ranging between a minimum of ~27% to a maximum of 100%. This means that, as already shown for LDA, the worst prediction can correctly attribute 3 out of 11 observations.

Figure 5.11 represents the trend of the prediction performance with increasing the number of test set objects.

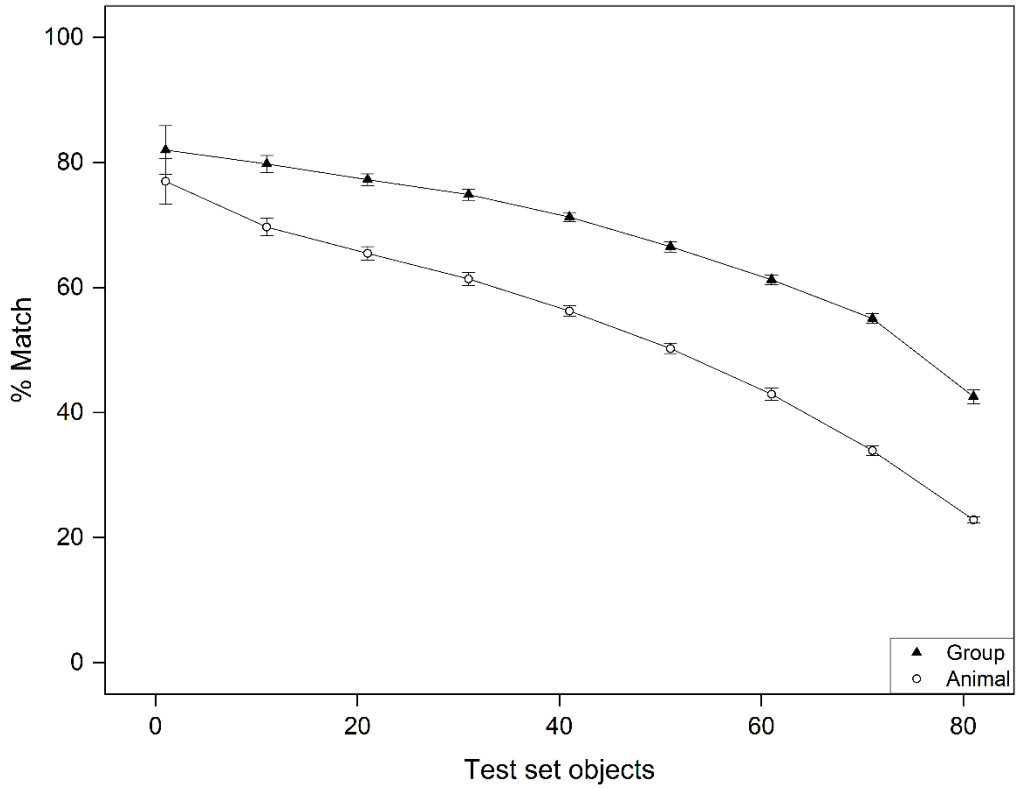

Figure 5.11. Percentage of match of the prediction of the classification tree model with increasing the number of objects of the test set. Triangles and dots represent the group and the animal classification, respectively.

Table 5.8. Match % of prediction changing the number of objects of the test set. Values are referred to the mean value of 100 cycles of the same model and 50 validation reiterations ± the standard deviation of the distribution.

Test set objects	Animal classification	Group classification
1	77 ± 4%	82 ± 4%
11	70 ± 1%	80 ± 1%
21	65 ± 1%	77 ± 1%
31	61 ± 1%	74.9 ± 0.9%
41	56.2 ± 0.8%	81.3 ± 0.7%
51	50.2 ± 0.8%	66.5 ± 0.8%
61	43 ± 1%	61.3 ± 0.8%
71	33.9 ± 0.8%	55.1 ± 0.8%
81	22.9 ± 0.5%	43 ± 1%

The shape of the curves reported in Figure 5.11 and the values reported in table 5.8 indicate a constant decline of the prediction performance with the decrease on the training set objects. The variability, expressed as the standard deviation, is almost constant except for the first points (77 ± 4% and 82 ± 4% in animal and group classification, respectively). For the group classification, the prediction match goes from 82 ± 4% to 43 ± 1% with a net percentage decrease of ~40 %, whereas for animal classification the performance goes from 77 ± 4% to 22.9 ± 0.5% with a decrease of

~54%. These percentage differences suggest that the classification tree model based on the group classification is more stable than the same model based on the animal classification.

5.4.3 Random Forests (RF)

Random forests (RF) were introduced by Leo Breiman (2001) and are one of the most used machine learning methods. The popularity of this method is due to its possibility to be successfully used in classification and regression problems, as well as in datasets with a small number of observations and/or a large number of variables (Biau and Scornet, 2016; Breiman, 2001).

RF are tree-based models involving multiple trees that are combined into a unique prediction model. The approach operates according to the "*divide et impera*" principle: each small portion of the data grows a tree model that is then aggregated to a larger predictive model (Biau and Scornet, 2016).

The difference between tree models and each tree that constitutes a RF is based on the number of predictors (variables) chosen for each split. Indeed, in RF a random choice of a limited number of predictors is used for each split. It has been reported that the most suitable number of predictors to be used for classification problems is $\sqrt{v}$, where v is the total number of variables (Hastie et al., 2021). This means that the RF model uses a random selection of variables whose number corresponds to approximately the square root of the total number of available variables. This approach was thought to overcome some problems related to the behavior of the variables in real datasets. Indeed, in a dataset where one or a few variables are strong predictors, together with moderately strong/weak predictors, the trees will be similar or even equal to each other because the algorithm would be prone to use just the most informative predictors (such as it happens in the tree). Pushing the algorithm toward the random choice of a subset of variables guarantees a larger variability of the model, including less informative variables somehow contributing to the description and the classification of the observations (Hastie et al., 2021). Hence, improving the model even with these less descriptive variables will influence the prediction power of the model.

5.4.3.1 Random Forests validation

RF model was performed using the CRAN package `randomForest`. The call was based on the default number of trees for the model (500) and the number of variables for each split was set to 3 as the total number of variables is 11 ($\sqrt{11} \cong 3$).

As seen for the validation of the previous models, the two levels of classification are reported: animals and groups. Table 5.9 reports the mean values and the standard deviation of the match percentages calculated on the match percentages of 50 reiterations of the same number of cycles. Figure 5.12 reports the prediction performance with the increasing of the model cycles from 1 to 1000.

Table 5.9. Mean and standard deviation of prediction match of RF model changing the classification parameter and the number of cycles.

Classification	1 cycle	10 cycles	100 cycles	1000 cycles
Animal	82 ± 14%	81 ± 4%	82 ± 1%	81.7 ± 0.3%
Group	90 ± 7%	90 ± 2%	89.9 ± 0.8%	89.7 ± 0.3%

Considering both the animal and group classification, the percentage of match is always above 80%, suggesting a greater predictive power with respect to the other models.

Even in the case of RF, the distribution with the highest variability is the model built on the animal classification with 1 repetition cycle (82 ± 14 %), showing a prediction percentage ranging between 45.5% and 100%. This indicates that in the worst prediction 6 out of 11 observations were mistakenly attributed to an animal class. In the previous models the worst performance accounted for 8 out of 11 attribution errors both for LDA and tree models.

Table 5.10 and Figure 5.13 report the trend of the prediction match with decreasing the number of objects in the training set from 95 (total number of observations = 96, minus 1 test observation) to 15 (96 total observations minus 81 test observations), with a step of 10. For each new training set, the model was operated at 100 rounds, and the reiteration of the validation was set to 50.

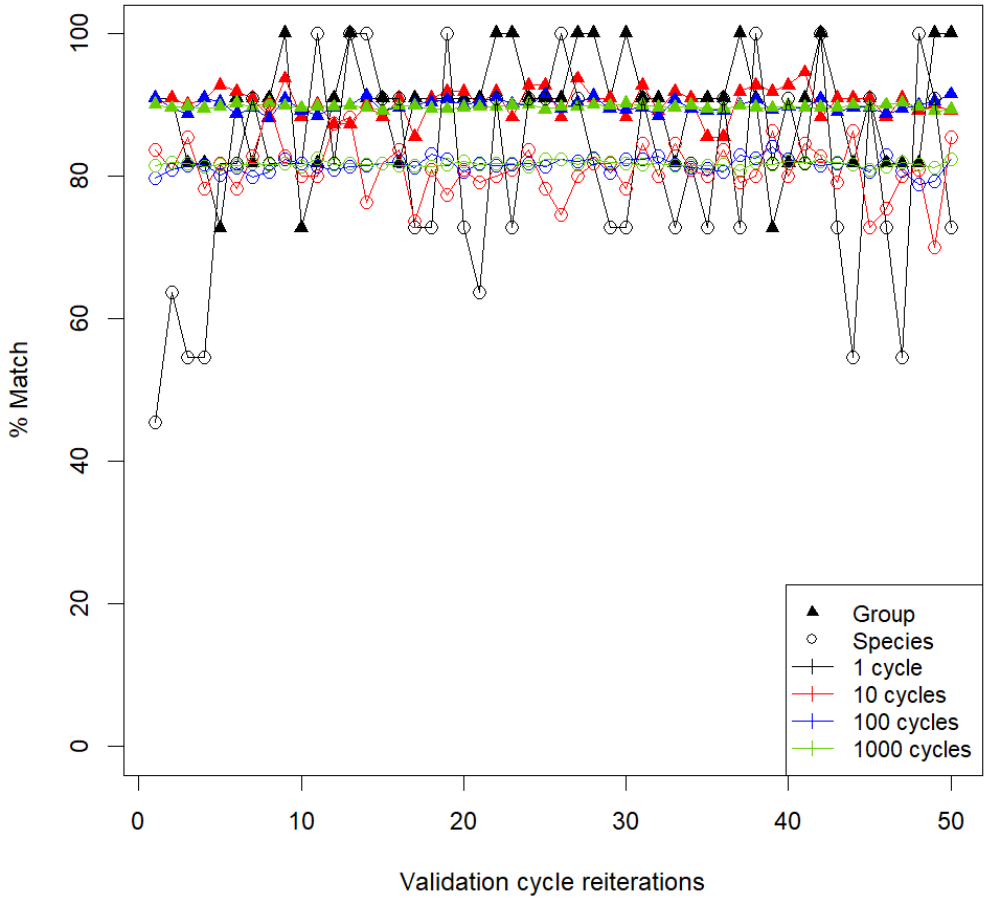

Figure 5.12. The percentage of match of the prediction using RF model varying the number of validation cycles (1, 10, 100, 1000). Empty dots are referred to the classification level "animal", whereas triangles indicate the classification level "group".

Table 5.10. Match % of prediction changing the number of objects of the test set. Values are referred to the mean value of 100 cycles of the same model and 50 validation reiterations ± the standard deviation of the distribution.

Test set objects	Animal classification	Group classification
1	85 ± 3%	91 ± 3%
11	82 ± 1%	89.7 ± 0.9%
21	78.9 ± 0.9%	88.7 ± 0.7%
31	75.7 ± 0.8%	87.6 ± 0.6%
41	72.5 ± 0.7%	86.1 ± 0.7%
51	68.5 ± 0.8%	83.8 ± 0.6%
61	63.9 ± 0.8%	80.6 ± 0.5%
71	57.4 ± 0.7%	75.7 ± 0.6%
81	48.3 ± 0.7%	67.6 ± 0.8%

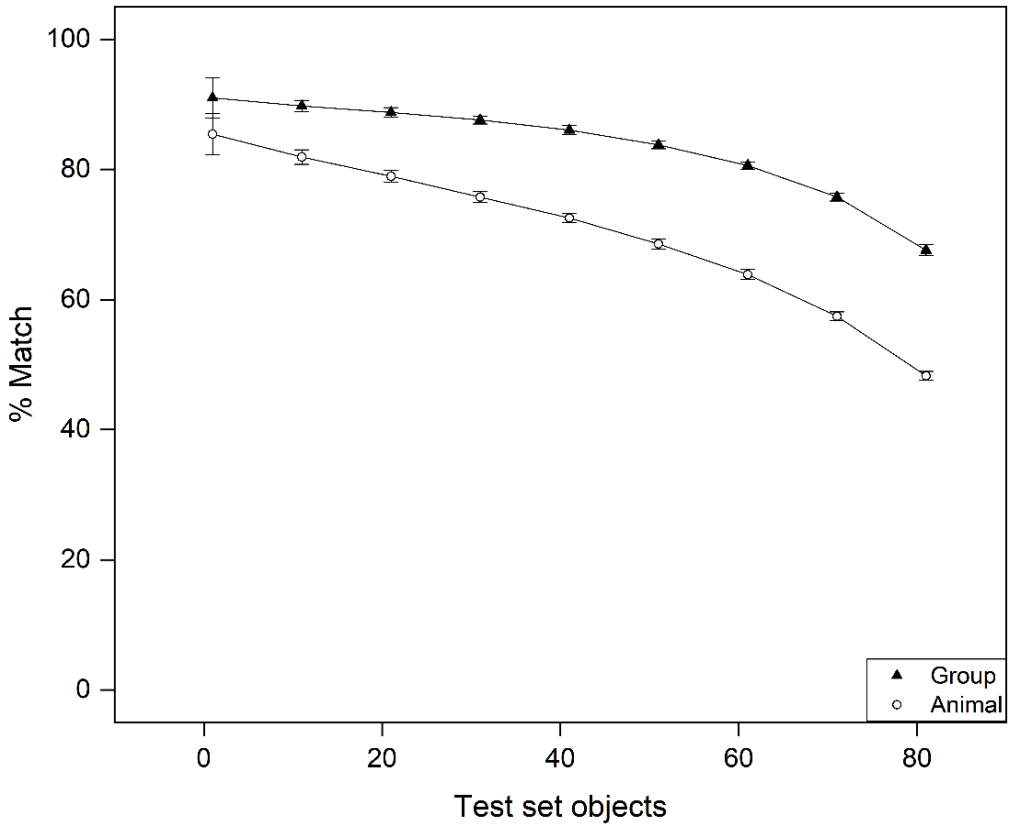

Figure 5.13. Percentage of match of the prediction of RF model varying the number of objects of the test set. Triangles and dots represent the group and the animal classification, respectively.

Looking at Figure 5.13 it is possible to see that the % match follow a constant decreasing trend for both curves. With the increase of the test set, the standard deviation of the distribution remains constant except for the first row, referring to the most numerous training set.

The net prediction performances determined for RF goes from 85 ± 4% to 48.3 ± 0.7% in the case of animal classification, and from 91 ± 3% to 67.6 ± 0.8% when the classification following the groups was performed. This implies a net decrease of ~37% and ~23% between the first and the last point of the curves of animals and groups, respectively, suggesting a greater robustness of the model when the group classification is performed, as also observed for the previous models.

5.4.4 Models comparison

The performances of the three models are compared to define which one performs better under the tested parameters. Table 5.11 reports the performances of the models as the classification and the number of cycles changes.

If compared by the animal classification level, the percentage of correct attribution is higher for RF (82 ± 14% for 1 cycle) with respect to both LDA and tree (64 ± 14% and 69 ± 14%, respectively, for 1 cycle). In addition, these last two models always present similar values when increasing the number of cycles. In a similar way, also the standard deviation shows the same behavior in all the models as the number of cycles increases from 1 to 1000.

Table 5.11. Match percentage for the two classification levels performed with the tree models increasing the number of cycles performed.

Classification	Model	1 cycle	10 cycles	100 cycles	1000 cycles
Animal	LDA	64 ± 14%	70 ± 4%	68 ± 1%	67.6 ± 0.5%
	Tree	69 ± 14%	69 ± 5%	70 ± 2%	69.9 ± 0.4%
	RF	82 ± 14%	81 ± 4%	82 ± 1%	81.7 ± 0.3%
Group	LDA	88 ± 8%	87 ± 2%	87 ± 1%	87.0 ± 0.3%
	Tree	80 ± 12%	80 ± 4%	80 ± 1%	79.7 ± 0.3%
	RF	90 ± 7%	90 ± 2%	89.9 ± 0.8%	89.7 ± 0.3%

Conversely, the prediction matches are different when classification groups are compared. Indeed, RF and LDA present similar values (88 ± 8% and 90 ± 7%, respectively, for 1 cycle), whereas the tree model shows the scarcest prediction accuracy (80 ± 12% for 1 cycle).

Additionally, the degree of variability with an increased number of cycles is comparable between the models, demonstrating that the variability is mostly influenced by the number of times that the models recalculate themselves, regardless of the type of model employed. Hence, the validation of the predictive power should be valued based on the number of times that the model recalculates itself, until a satisfying variability is achieved.

Figure 5.14 represents the prediction match with the shrinking of the training set. The shapes of the distributions allow to identify different behaviors of the models: LDA curves (in blue and light blue) are characterized by an almost constant value until a point where the slope increases, whereas classification trees and RF (red and green curves) display more constant slopes on the whole interval. This could be related to the

similar mechanism of algorithm construction: RF models are based on the construction of an agglomeration of trees, as reported in paragraph 5.4.3.

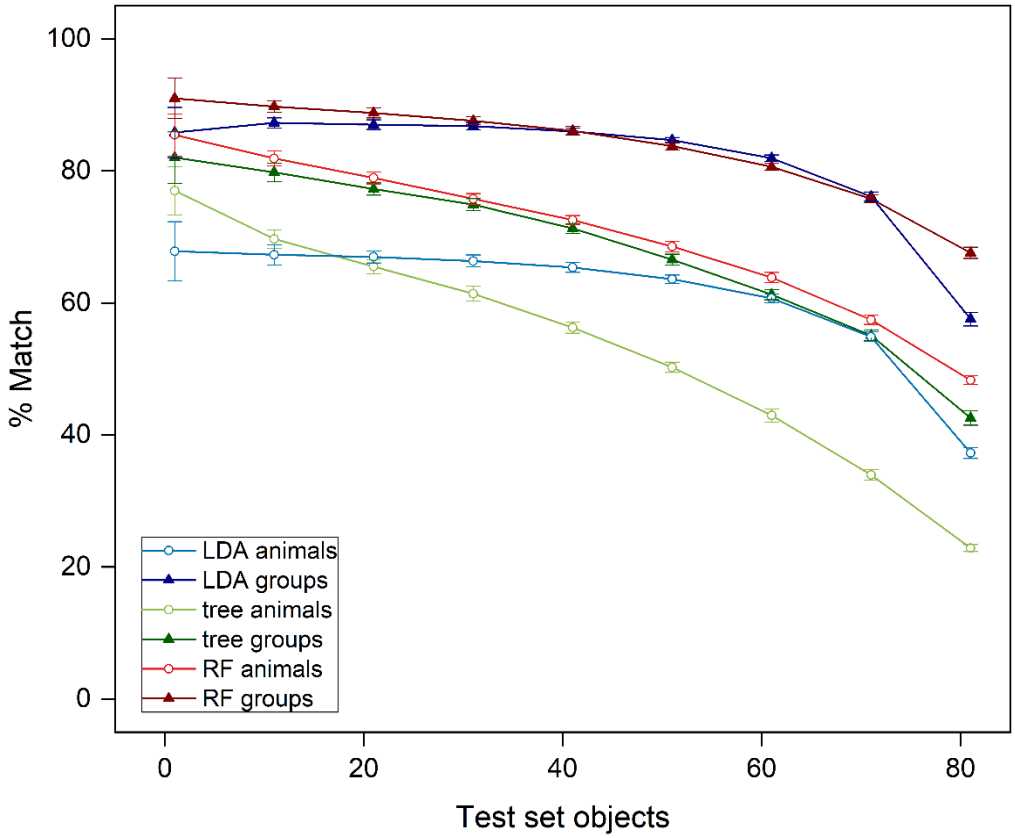

Figure 5.14. Percentage prediction match of the models with the change of the number of objects of the test set. Triangles and dots represent the group and the animal classification, respectively. Blue curves are referred to as LDA, green curves to tree, and red curves to RF.

Looking at the curves of the animal classification level (empty dots), the highest match percentage is the one referred to the RF model, followed by trees, and lastly LDA. The decreasing trend continues until the last point of the test is reached, for which the best performance is once again shown by RF, followed by LDA, and lastly by the tree model.

When the curves of the classification level groups are compared, the percentages are slightly different: in correspondence with the first and the last points of the curves, the prediction rank of the models is RF > tree > LDA. In the central part of the curves, similar % matches between LDA and RF are observed.

5.4.5 Mixtures prediction

For this series of tests, the prediction performances of the models were tested on a dataset where the fecal sterols and stanols were calculated to simulate mixtures of animal and human feces.

The fecal sterols and stanols from the custom mixtures were used as test set for the models. This dataset aimed to simulate the result expected when a mixed input source of animals and/or humans is expected. All the possible groups permutations of the following binary ratios were included in the dataset: 90:10, 80:20, 70:30, 60:40. The dataset was constructed using a random choice of 60 observations for each composition and for each possible class permutation. The dataset is composed of 2400 objects resulting from the random choice of two objects and mixed following the set ratios.

The models were built on the dataset obtained from previous experiments (i.e., the dataset resulted from the analysis of fecal sterols and stanols for equines, caprine, swine, bovines, and humans).

The classification level decided was groups, for the following reasons:

a. The accuracy of the prediction is greater for group classification rather than for animal classification.
b. The models based on the group classification are more robust.
c. The genera or subfamily level of detail is acceptable enough in the research of historical/archaeological interest.

The results of the match performance were evaluated on the correct prediction of both the animals (from now on reported as total combination-TC) in the mixture, the most present (from now on reported as primary component-PC), and the less present constituent in the mixture (hereafter reported as secondary component-SC). The results are reported in Table 5.12.

Table 5.12. Percentage of match of the tree models when the mixed inputs are tested. The first column reports the total combination correctness, while the following columns refer to the %match of PC and SC.

	TC match %	PC match %	SC match %
LDA	34%	85%	35%
Tree	20%	73%	24%
RF	30%	81%	32%

The percentage of the cumulative correct prediction, corresponding to the perfect estimate of both the contributors to the mixture, is higher in LDA (34%), followed by

RF (30%), and tree (20%). The percentages of correctness are greater for all the models when only one contributor is evaluated: LDA is able to determine the PC in the 85% of the tested observations, followed by RF and lastly tree, with accurate predictions in the 81% and 7 % of the cases, respectively. The percentage of match of the SC is lower but follows the same order: LDA 35% > RF 32% > tree 24%.

From these results it appears that the model exhibiting the best prediction power, from the perspective of the global correctness of the predictions, is LDA.

To investigate the predictive power of the models more in depth, the percentages of correct prediction when changing the ratio of the mixtures were calculated (Table 5.13).

Table 5.13 Percentages of match of the models as the mixture proportion varies, TC, PC and SC are reported.

	Model	90:10	80:20	70:30	60:40
TC match %	LDA	26%	33%	40%	38%
	Tree	22%	22%	23%	15%
	RF	29%	33%	32%	28%
PC match %	LDA	93%	91%	86%	68%
	Tree	92%	82%	68%	51%
	RF	98%	91%	76%	58%
SC match %	LDA	37%	34%	40%	40%
	Tree	23%	28%	28%	21%
	RF	29%	33%	33%	31%

By changing the ratio of the two components of the mixtures, the performances of the models follow different pathways. For LDA, the TC match is higher when the mixture 70:30 is tested (40 % of correctness of both the components of the mixture) and a similar value, 38 %, is observed for the ratio 60:40. Lower percentages of accuracy are calculated when the ratios 90:10 and 80:20 are examined. On the other hand, when the prediction results of the tree model are analyzed, the percentage of match remains almost constant when modifying the mix ratio from 90:10 (22%) to 80:20 (22%) to 70:30 (23%) and decreases for the 60:40 (15%) ratio. In turn, RF presents almost the same prediction match with the change of the mix proportions: 29% for the mix 90:10, 33% for 80:20, 32% for 70:30, and 28% for the ratio 60:40.

Moving to the percentage of correctness of the PC, a completely different scenario is observable: the percentage of match is indeed always above 50%, with peaks of 98 % for RF when the ratio 90:10 is tested. In this case, the model match percentages follow the order RF > LDA > tree, going from 98% to 92% of correctness. For the 80:20 proportion LDA and RF present the same match percentage (91%), whereas tree shows a lower value of prediction match (82%). The percentage of the correct prediction when the two proportions 70:30 and 60:40 are observed follows the order LDA > RF > tree

providing values of 86%, 76%, 68% and 68%, 58%, 51%, for the two proportions and three models, respectively.

The prediction matches of the SC are higher for LDA for all the mix ratios and span between 34% (for the ratio 80:20) and 40% (when the proportions 70:30 and 60:40 are analyzed), followed by RF with almost constant values of about 30%; lastly, the tree model predicts correctly the SC with values between 21% (60:40) and 28% (for both 80:20 and 70:30).

The results of the match accuracy of the PC shows, intuitively, that best performances are correlated to the higher proportion of the PC itself: the more of one component is present in the mixture, the higher will the predictive power be, regardless of the model performed. This behavior is not always reflected by the performances of the correct prediction of the SC: different trends are observable for the three models. LDA seems to distinguish the SC in both higher (mix ratios 70:30 and 60:40), and lower concentrations (38% match for the 90:10 proportion). Both in the case of the tree model and of RF, it is more difficult to define a precise trend as both present the best match predictions for the 80:20 and 70:30 proportions. Possibly, a lower ambiguity of the algorithms is achieved when the SC concentration is sufficiently high to be tagged (typically 20-30%), but not so much as to be 'blended' with PC (ratio 60:40). These predictive patterns are reflected by the combined contribution match (TC).

The second question raised from the observation of these results was if there is a group of animals that is more prone to be mistakenly predicted. In other words, is (are) there a group (groups) for which the error in the prediction is greater?

The percentage of accuracy in the prediction for the group classification was determined and reported in Table 5.14 and Figure 5.15.

Table 5.14. Percentage of correctness in the prediction of the animal groups divided by PC and SC, and by model.

	Model	Equine	Bovine	Caprine	Swine	Homo
PC match %	LDA	82%	76%	87%	85%	95%
	Tree	63%	75%	79%	84%	65%
	RF	76%	88%	85%	90%	67%
SC match %	LDA	42%	27%	33%	36%	38%
	Tree	31%	17%	26%	25%	25%
	RF	33%	37%	27%	36%	30%

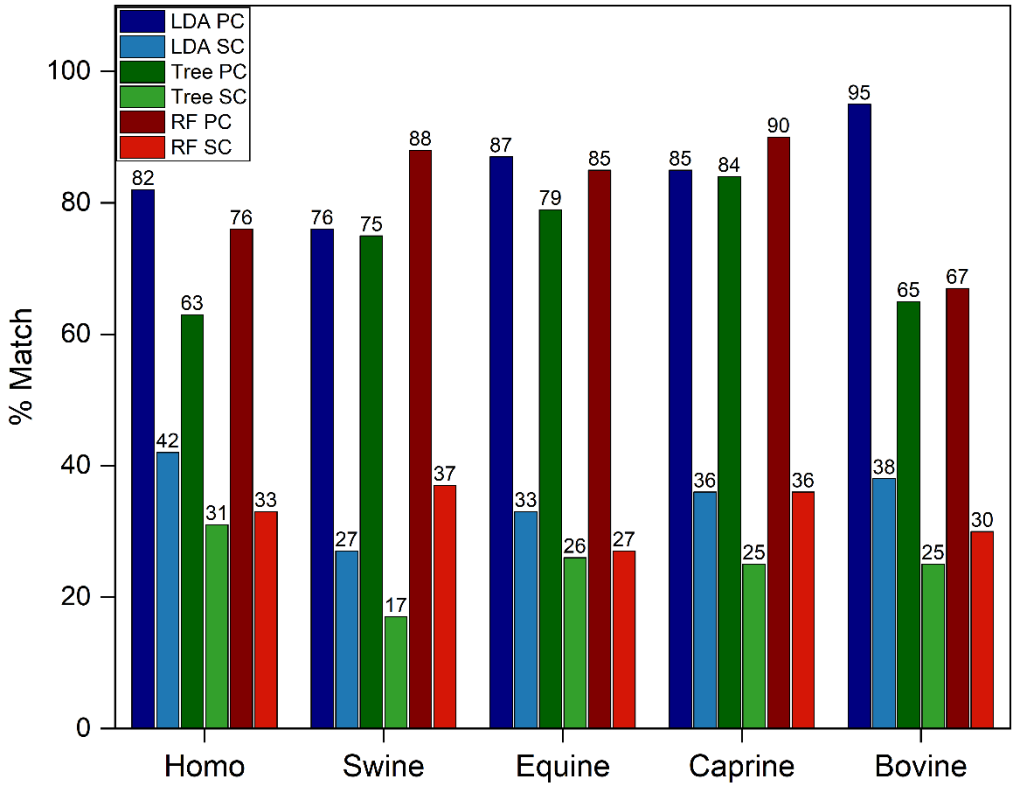

Figure 5.15. Percentage of match for the mixed observations divided by group. Different colors correspond to the model: blue for LDA, green for tree, red for RF. Dark shade of each color is referred to the PC, light shade indicates the SC.

The columns showing the PC match % obtained through LDA always have values above 75%, meaning that three quarters of the tested objects were correctly classified for the PC of the mixture. As displayed in Figure 5.15, for LDA the group with the greatest percentage of match on the PC is bovine (95%) followed by equine (87%), caprine (85%), homo (82%), and swine (76%); whereas a different performance order is determined if looking at the SC: homo > bovine > caprine > equine > swine, with percentage match values spanning between 27% and 42%.

When the performances of the tree model are analyzed, the percentages of match for the PC span between a minimum of 63% and a maximum of 84%. The lowest value is observed for homo, whereas the best prediction performance was found for the 'caprine' group, followed by equine (79%), swine (75%), and bovine (65%). The predictions of the SC are the worst of the three models tested, ranging between 17 % and 31%. This means that in the best case (homo), the SC is correctly classified in less than one third of the observations. The animal group with the worst prediction match is swine, slightly higher percentages are observed for bovine and caprine (25%), followed by equine (26%).

RF produced better match predictions than the tree model both on the PC and SC. The group with the greatest match correctness is caprine when it is present as PC (90%), followed by swine (88%), equine (85%), homo (76%), and lastly bovine (67%). Slightly different order of prediction accuracy is observed if looking at the SC: swine is the group with the best prediction accuracy (37%), followed by caprine (36%), homo (33%) bovine (30%), and equine (27%).
Appendix E reports all the match percentages for all the possible permutations of the mixed combinations.

From the results reported in this paragraph, it is possible to define LDA as the model showing globally the best performances in the predicting components and in their matching. Looking at the results of the animal groups, LDA is the model that generally performs best, even if it is not the best if the single animal groups are analyzed.

5.5 Fecal sterols in Dark Earth

The aim of the evaluation of the prediction parameters was to attempt a qualitative classification of the samples coming from the sites of via Pigna (VP), vicolo San Pietro in Monastero (SPM), and Castelar di Leppia (CDL).
Considering the results of the tests on the models shown in the previous sections, the best prediction performances were obtained for LDA, hence considered the best candidate for the prevision of the determination of the fecal content in the case of Dark Earth (DE) samples.
The number of samples analyzed for each profile and sites are reported in Table 4.1.
Before moving into a brief description of the sterols distributions, the results of the ratio COP/CHN are reported. This ratio was proposed by Leeming et al., (1996) to determine whether the soil was polluted by feces or not. The results (expressed as mean value ± standard deviation) obtained for each site are SPM: 3.0 ± 0.8, VP: 2.9 ± 0.9, and CDL: 0.5 ± 0.1. Values exceeding 0.5 indicate the presence of fecal pollution, otherwise the contamination is considered negligible.
Both VP and SPM resulted to be highly contaminated by feces, whereas CDL shows lower values. Despite this, coprostanol is not always the most present congener in

animal feces (section 5.1) that, hence, can account for a certain input in CDL, but only a relatively small presence of human (or omnivore) feces, on average constituted by ca. 60% of coprostanol (Leeming et al., 1996), can be supposed. The distributions of sterols and stanols are reported in Table 5.13 and in Figure 5.17. The bars represent the mean value of the compound in the distribution of samples, the error bars represent the confidence interval (95%).

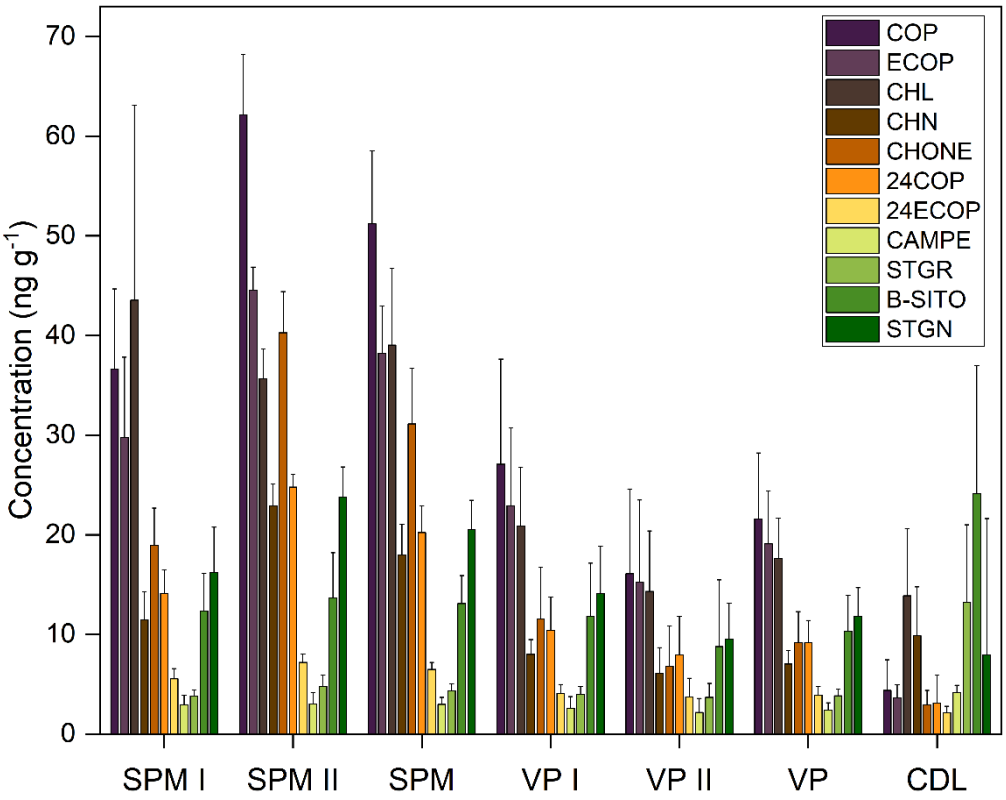

Figure 5.17 Sterols and stanols distribution for the profiles and sites SPM I, SPM II, SPM, VP I, VP II, VP, and CDL. Columns represent the mean values for each analyte, error bars are referred to the 95% confidence interval of the distribution.

Table 5.13. Concentration of sterols and stanols for profiles and sites values are expressed as mean value ± 95% confidence interval of the concentration in ng g^{-1}.

Profile/Site	COP	ECOP	CHL	CHN	CHONE	24COP	24ECOP	CAMPE	STGR	β-SITO	STGN
SPM I	37 ± 8	30 ± 8	45 ± 20	12 ± 3	19 ± 4	14 ± 2	6 ± 1	3 ± 1	3.8 ± 0.6	12 ± 4	16 ± 4
SPM II	62 ± 6	45 ± 2	36 ± 3	23 ± 2	40 ± 4	25 ± 1	7.2 ± 0.8	3 ± 1	5 ± 1	14 ± 4	24 ± 3
SPM	50 ± 7	38 ± 5	39 ± 7	18 ± 3	31 ± 5	20 ± 3	6.5 ± 0.7	3 ± 1	4.3 ± 0.7	12 ± 3	20 ± 3
VP I	27 ± 10	23 ± 8	21 ± 6	8 ± 1	11 ± 5	10 ± 3	4.1 ± 0.9	3 ± 1	4 ± 1	12 ± 5	14 ± 5
VP II	16 ± 8	15 ± 8	14 ± 6	6 ± 3	7 ± 4	8 ± 4	4 ± 2	2 ± 1	4 ± 1	9 ± 7	9 ± 4
VP	22 ± 7	19 ± 5	18 ± 4	7 ± 1	9 ± 3	9 ± 2	4 ± 1	2.4 ± 0.7	3.8 ± 0.7	10 ± 3	12 ± 3
CDL	4 ± 3	4 ± 1	14 ± 7	10 ± 5	3 ± 1	3 ± 3	2.1 ± 0.6	4.1 ± 0.7	13 ± 8	24 ± 13	8 ± 14

In SPM I the three most present congeners are CHL (45 ± 20 ng g^{-1}), followed by COP (37 ± 8 ng g^{-1}) and ECOP (30 ± 8 ng g^{-1}). In SPM II, the most concentrated congeners are COP (62 ± 6 ng g^{-1}) > ECOP (45 ± 2 ng g^{-1}) > CHONE (40 ± 4 ng g^{-1}). Cumulatively, the site SPM is enriched in COP (50 ± 7 ng g^{-1}), CHL (37 ± 7 ng g^{-1}) and ECOP (38 ± 5 ng g^{-1})

In VP I and VP II the highest bars are the ones referred to COP (respectively: 27 ± 10 ng g^{-1}, 16 ± 8 ng g^{-1}), ECOP (respectively: 23 ± 8 ng g^{-1}, 15 ± 8 ng g^{-1}), and CHL (respectively: 21 ± 6 ng g^{-1}, 14 ± 6 ng g^{-1}). For the entire site VP, the prevalent congeners are COP (22 ± 7 ng g^{-1}), ECOP (19 ± 5 ng g^{-1}) and CHL (18± 4 ng g^{-1}).

In CDL the major contribution given by β-SITO (24 ± 13 ng g^{-1}), followed by CHL (14 ± 7 ng g^{-1}) and STGR (13 ± 8 ng g^{-1}).

In the profiles of SPM and VP, a generally higher contribution given by cholesterol-derived compounds (than by CDL) can be appreciated. On the other hand, CDL displays the most relevant input of plant-derived sterols and stanols. These diverse distributions could be associated with a different fecal input in the soil: feces from animals following an omnivore diet are expected for VP and SPM, whereas CDL would be likely more influenced by the contribution of herbivores and plants. These abundances are partially in accordance with the results from the micromorphological analysis (Appendix B), showing variable amounts of fecal material in SPM and VP, and the almost complete absence in CDL.

The dataset of the concentrations of sterols and stanols in DE samples was used as test set for the prediction when applying the LDA classification model previously built on the content of sterols and stanols of the animal and human feces. Prior to the application of the prediction function, the observations were normalized to the total sum of sterols and stanols for each sample, in order to reduce the variability and to perform the same data treatment employed for the animal feces dataset. The results of the prediction generated applying LDA to the DE samples are reported in Table 5.14.

Table 5.14. Cumulative percentage of predicted groups of classification divided by profiles and sites.

	Homo	Swine	Equine	Caprine	Bovine
SPM I	0%	100%	0%	0%	0%
SPM II	0%	92%	0%	0%	8%
SPM	0%	95%	0%	0%	5%
VP I	0%	83%	0%	0%	17%
VP II	0%	100%	0%	0%	0%
VP	0%	92%	0%	0%	8%
CDL	0%	0%	66%	0%	33%

The results represent the group of animals associated to the highest probability. The percentages reported in Table 5.14 refer to the number of times that a specific class is predicted on the total number of observations for each profile and site. For example, in SPM I, the totality of the samples is classified as 'swine', whereas in SPM II a small set of samples is classified as bovine (8%) while the majority is predicted as swine (92%).

In a similar way, the most common class in VP I is swine (83%), but a little portion of the observations (17%) is associated with bovine. The samples of profile VP II are all classified as swine. Hence, for the whole site VP, most of the contribution would be given by swine (92%), followed by bovine (8%).

In CDL observations are mostly classified as equine (66%) and the remaining 33% as bovine.

As reported previously when discussing the distributions (Figure 5.17), a dominant presence of omnivores (swine) or of animals with a higher cholesterol-derived compounds in their feces (bovines) have been associated to VP and SPM, whereas CDL predominant input consists of plant derived sterols. The latter can be closely related to the presence of animals with a strictly herbivorous diet, as donkeys and horses.

Despite none of the observations was classified as 'homo', the strong anthropogenic presence in the sites cannot be excluded; indeed, the presence of humans in the sites has been demonstrated by the proxies presented in the previous sections (Chapters 3 and 4). In addition, the prediction performances of the LDA model, the best among the tested ones, are always related to a percentage of incorrect classification. Moreover, regardless of the models being tested on a dataset where the mixture of fecal input was calculated, they were built under one level of classification into separate groups. In this way, the regions of uncertainty of the classification models are significant if different fecal input contribute to the nature of the samples (Hart et al., 2000). Under these conditions, such samples would fall into the areas close to the boundary of a hyperspace plane, forcing the model to a unique classification, and not reflecting the complex nature of the sample. Therefore, the classification models based on the probability tested in this chapter do not fully resolve the question about the fecal contribution in Dark Earth samples. Thus, the need for specific classification models able to identify the presence of different inputs and, possibly, return a concentration of the sources is evident when the coexistence of humans and animals in the same site can be assumed.

Conclusions

This work considered DE excavated in three archaeological sites located in the Verona city center (vicolo San Pietro in Monastero and via Pigna) and in the historical village of Castelar di Leppia. Samples were investigated through a multidisciplinary and multi-technique approach. This conclusive section is divided into different paragraphs related to the main contents discussed in the previous chapters.

Dark Earth formation: human activities and soil exploitation

From the observation of the thin sections, common features and different syn- and post-depositional phenomena of DE from three archaeological sites were observed and it was possible to identify the fast-growing processes responsible for the formation of studied the DE layers. Particularly, the dumping of heterogeneous material combined with the reworking of older sediments can be considered as their basic formation mechanism. In detail, at SPM the dumping of heterogeneous material, where older sedimentary matrices and older components are mixed with inclusions dating closer to the time of formation, seems the most probable setting process. Indeed, the occurrence of hiatuses, and few signs of bioturbation suggest a fast-growing mechanism in both SPM and VP. On the other hand, the more expressed bioturbation process observed in CDL indicates a slower rate of accumulation of the soil.

The fast-growing hypothesis as the result of dumping of heterogeneous material is also supported by the different radiocarbon dating of humin fraction and charcoal inclusions, with the former being older. Thus, an interesting aspect resulting from the radiocarbon data is the possibility to obtain different information by choosing different organic fractions in cultural layers: the date of the humin fraction suggests the first dwelling of the site, while charcoal dating could reveal the period of soil exploitation.

The anthropogenic influence on DE soil was underlined and it was possible to highlight the specificity of some features. The material used for backfilling (or dumping) is likely composed of organic material and charred wooden particles, coming from manufacturing activities. Therefore, the differences between sites were evident.

From the distribution of charcoal fragments, trace elements and organic/inorganic carbon it is possible to highlight similarities within the profiles of the same sites (i.e., SMP I-SMP II and VP I-VP II), and that neither trends nor internal differentiations can

be found. This occurrence suggests that the studied layers have likely been the result of different processes of soil exploitation (backfilling and dumping of material mixed with the matrix and sometimes reworked), in agreement with the absence of chronological distribution of the radiocarbon dates.

The significant differences observed between the sites for the different carbon fractions, particularly between SPM and VP, suggest that the horizontal variability in DE, likely related to different practices such as agriculture or waste management, is significant even for short-range distances. In particular, charcoal fingerprints of the three sites differ even when there is proximity in terms of age and location, this proves the occurrence of a horizontal rather than vertical variability of the DE. This aspect suggests a strong and diversified human influence on the studied DE. Indeed, VP and SPM show differences also in terms of TOC, ROC, and charcoal distributions, although these two sites belong to DE deposits that can be correlated.

The human influence in terms of use of the soil for domestic purposes is further marked by the presence of metallic droplets and trace element concentration (i.e., P, Fe). This points to a process whereby humans affected the natural soil both in VP and SPM. The site VP seems to be situated in the immediate vicinity of artisanal activities (e.g., metal, or ceramic manufacturing), while SPM seems to be more likely affected by the gathering of remains from the close by VP site. High P content in all the sites agrees with the presence of anthropogenic influence (Migliavacca et al., 2013), as the presence of P has already linked to human activities, being associated to excrement and byre addition and/or to the dumping of biomass burning residues (Devos, 2018). In this case, the excess of P in VP might be related to this latter activity, in agreement also with the abundant presence of charred fragments and ROC in this site with respect to SPM and CDL.

Fire occurrance

The differences in charcoal distribution observed for the tree sites may be associated to distinct dynamics of production and/or transport of the biomass burning residues. This might in turn be related to independent anthropogenic activities that occurred in each site, such as the metallurgical activities described previously for VP. Additionally, the presence of fine biomass burning residues in the samples of CDL is explained by the prolonged time of bioturbation (CDL deposit is ca 1500-1800 years older with respect to VP and SPM), supported by the thin sections analysis.

Raman results obtained on the charcoal particles showed that fragments from CDL are characterized by a higher level of structural disorder with respect to the charred particles selected for the SPM and VP sites.

This finding could be explained by the variation in age, as CDL is likely older than the others. Being older and probably more degraded, charcoal had more disordered structures with a lower fraction of graphitized carbon (Smidt et al., 2020, 2017). This aging process may be ascribed to environmental factors, such as pH and humidity of the soil, microbial activities or water erosion (Arroyo-Kalin, 2020; Inoue et al., 2017; Rumpel et al., 2015). Interactions of charred residues in soils due to chemical and physical changes have been investigated and suggested that the main changing features of charcoal could be due to physical and biological erosion in soils (Cheng et al., 2008; Inoue et al., 2017). Moreover, data demonstrated the occurrence of lower burning temperatures for the charcoal of CDL. Such evidence implies that the more ancient charcoal of CDL was produced at lower carbonization temperatures, providing a limited removal of amorphous structures and a scarce formation of graphene-like networks (McDonald-Wharry et al., 2013). On the other hand, the similarity between the burned particles from SPM and VP supports the assumption that SPM and VP are coeval, while higher carbonization temperatures may be related to high-energy demand processes, such as the manufacturing activities.

The correlation between the results obtained from different techniques revealed the relation between the most recalcitrant fraction of carbon (ROC) measured in the soil samples and the charred particles collected into the DE samples. But the absence of significant correlations with PAHs and Levoglucosan could indicate a different mechanism of formation of ROC or a different outcome in the environment. Indeed, ROC has been identified as the carbon pool with the lower rate of bio-adsorption in soil, and it has been associated to complex and high weight class of molecules such as humic acids (Kleber, 2010) and black carbon in the form of benzenepolycarboxylic acids (Llorente et al., 2010). In this framework, PAHs and levoglucosan would not be included in the ROC fraction and could be also more prone to biodegradation.

In the three sites investigated, some differences have been noticed both in terms of content of different classes of weight of PAHs and to the proxies for the identification of the PAH sources. It was found that SPM is the site with higher enrichment of PAH, followed by CDL and VP. The source of PAH is mostly pyrogenic, and it was possible to identify PAHs derived from softwood combustion in SPM and VP, whereas in CDL

the PAHs distribution accounts for different sources of biomasses, namely hardwood and/or grasses and shrubs.

Plant input and animal presence

The environmental conditions were described using the *n*-alkanes: the main contribution of alkanes seems to be related to terrestrial plants due to the strong abundance of heavier *n*-alkanes and the values of the ACL in the three sites. The different *n*-alkanes profiles of VP, SPM and CDL, suggest that trees were more present in VP and SPM, whereas grasses and shrubs were likely more present in CDL. This finding is also supported by the concentration of miliacin detected in the three sites, with a slight abundance of this biomarker in CDL. Possibly, this difference in vegetation reflects the different biomasses burned in CDL, with respect to VP and SPM. The degree of degradation of the hydrocarbons, expressed by the OEP, points to different soil conditions: the higher degree of conversion or the *n*-alkanes in VP and SPM, is possibly imputable to a periodic oxygenation due to a remixing of the soil, whereas CDL seem to be less affected by biodegradation phenomena, indicative of a lower level of soil manipulation. Additionally, the enrichment of VP and SPM in electron acceptors elements, such as Fe, could have promoted the degradation of the *n*-alkanes by soil microbiota.

From the analysis of fecal sterols and stanols, it was possible to identify a strong influence of fecal pollution, suggesting that these soil layers would be considered the result of the mixing of heterogeneous materials. Specifically, results demonstrated the presence of omnivores feces in the highly anthropized areas of VP and SPM. Conversely, CDL samples showed pollution mostly coming from herbivorous animals. The presence of mixed fecal input in the studied samples is probable, but as it stands, and with the classification models explored in this thesis, it is impossible to derive a numeric quantification attributable to different inputs.

By means of all this evidence it is possible to conclude that the information inferred from the DE cultural layers is extremely local and associated with a sequence of human activities and soil processes specific to a particular succession of events, even when the sites are close together. Hence, the presence of human activities for centuries on the sites of interest could have promoted a process of input of external material and mixing with the original matrix, forcing the system to rearrange and re-equilibrate. The strong anthropogenic impact is evident from the reconstructions of the proxies from the sites

of VP and SPM. These sites are influenced by a combination of activities, that results in a complex series of sources that have influenced these layers. Conversely, the site of CDL seems to be less affected by human pressure, where agricultural activities could have concurred to the DE formation.

In conclusion, the term 'dark earth' is just a label, under which the complex mechanisms of formation and development of the soil fall. These are strictly correlated to the human daily needs, which pervade all the surrounding space in such a deep manner, that it can resist through the centuries and arrive to us. Hence, even if the soil layers from the different sites appear to be similar from a macroscopic point of view, the differences in the content of the proxies investigated and the microscopic investigation suggest that several mechanisms brought to the cultural layers settling which appears to be unique to a specific area and time period.

Further perspectives

This doctoral project wanted to investigate the possibility of a multi-analytical approach in the context of archeological research. The perspective of a multidisciplinary protocol in this field of study could be helpful, beyond the sites studied in this thesis. In fact, from the results shown in this work, the integration of traditional investigation techniques and new analytical methods should be followed, pursuing a deeper comprehension of the mechanisms that influenced the environment since the beginning of human culture. Notably, the integration of the global results from the micromorphological analysis integrated with the specificity of the chemical biomarkers appears to be a valuable tool for the reconstruction of events and domestic habits in the past.

Moreover, the Raman spectroscopy appears particularly promising for the reconstruction of past fire conditions analyzing charcoal. Thus, the relative cost efficiency of Raman spectrometers makes this technique accessible and potentially interesting for the determination and characterization of soil inclusions.

Further perspectives regard implementing a model able to distinguish and quantify the contribution of different sources of proxies starting from separate group classifications. In particular, the interest will be put on the determination of the presence of animals and humans in a site starting from the analysis of fecal input. This concern, indeed, regards not only the case of dark earth, and more generally the cultural layers, but all the possible sites where the combined presence of humans and animals is supposed.

Appendix A

Analysis of sterols and fragmentation

This appendix is a supplementary analysis on the fragmentation pattern of the sterols and stanols. The following considerations are referred to the molecular weights of the TMS derivative compounds reported in the table A.1. In the following paragraphs the formula [M-X] will be used to specify the fragments, where M is the molecular weight and X is the number of the fragment lost during the ionization. The carbon numbers are indicated with the *Ci* abbreviation where *i* is referred to the position of the carbon atom following the numeration reported in figure A.1.

For Δ^5-sterols, 5α-stanols and 5β-stanols and ion with the ratio mass/charge (m/z) $[M-90]^+$ is common. This fragment represents for the loss of the trimethylsilanol (TMSiO) group and $[M-73]^+$, where the m/z 73 corresponds to the ion [TMSi]; another common ion is the one related to the loss of a methylic group in C19 position resulting in m/z $[M-15]^+$ (Goad and Akihisa, 1997; Partridge and Djerassi, 1977). For Δ^5-sterols the ion with m/z 129 is commonly present in the mass spectra, this ion is related to the loss of the trimethylsilane ester with the C1, C2, and C3 from the A ring (Brooks et al., 1968; Diekman and Djerassi, 1967; Goad and Akihisa, 1997). FigureA.2 reports the fragmentation pattern described for a generic sterol molecule.

Table A.1. Names, classification, and molecular weights (MW) for sterols and stanols used in this research. MW according to NIST database except for (*) marked that comes from Pubchem database.

Trivial name	Abbreviation	Classification	MW	MW TMS derivative
Cholesterol	CHL	Δ^5-sterol	386.6535	458.8347
Cholestanol	CHN	5α-stanol	388.6694	460.8505
Coprostanol	COP	5β-stanol	388.6694	460.8505
Epi-coprostanol	ECOP	5β-stanol	388.6694	460.8505
Campesterol	CAMPE	Δ^5-sterol	400.6801	472.8612
β-sitosterol	β-SITO	Δ^5-sterol	414.7067	486.8878
Stigmastanol	STGN	5α-stanol	416.7226	488.9037
Stigmasterol	STGR	Δ^5-sterol	412.6908	484.8719
24-ethyl-coprostanol	24-COP	5β-stanol	416.72(*)	488.9011
24-ethyl-epi-coprostanol	24-ECOP	5β-stanol	416.72(*)	488.9011
Cholestanone	CHONE	stanone	486.7 (*)	-

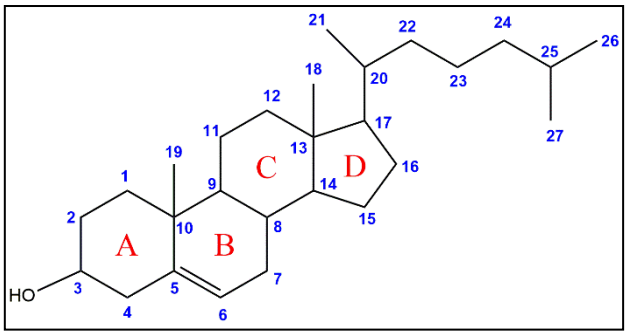

Figure A.1. Structure of cholesterol molecule with the carbon numbers following Goat and Akihisa (1997).

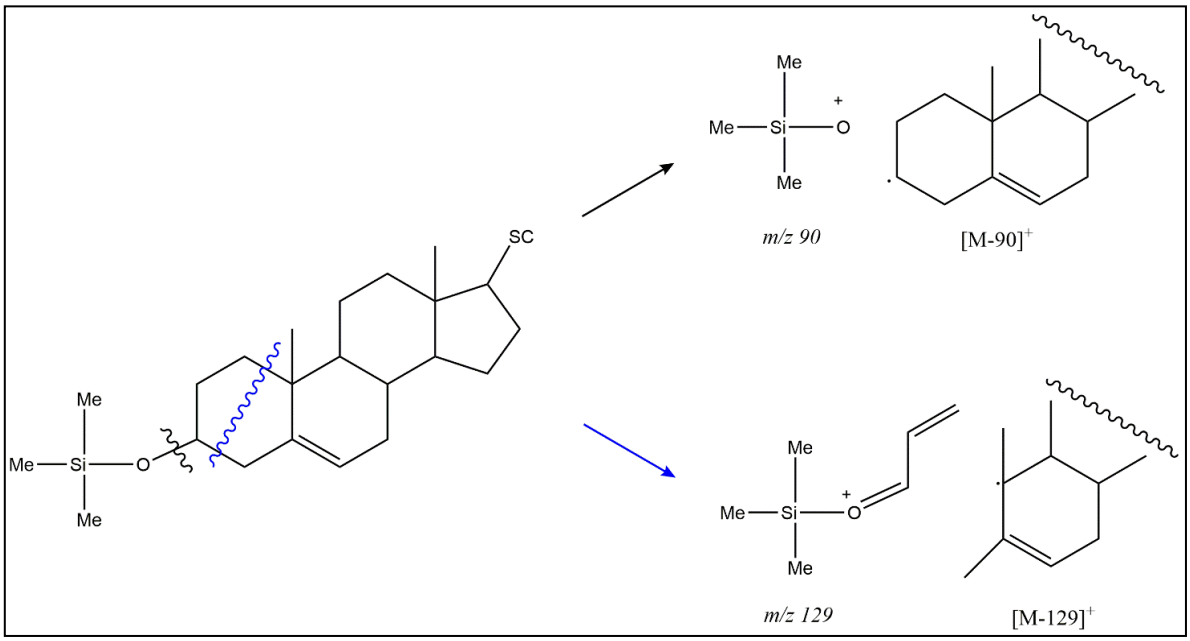

Figure A.2. Fragmentation pattern and formation of the common molecular ions with m/z 90 and 129.

In the mass spectra of both 5α and 5β stanols the signal of the TMS-esters of the molecular ion is present, in addition, 5β-stanols can give the $[M-32]^+$ ion for the loss of methanol (Goad and Akihisa, 1997; Idler et al., 1970).

The loss of the side chain and part of the D ring (C15, C16 and C17, with 42 a.m.mu) is common in many sterols and stanols, the main fragments involved in this mechanism are reported in Figure A.3. This fragmentation is coupled with the loss of the TMS-ester, the resulting ion has m/z $[M-side\ chain-42-TMSE]^+$. The remaining big fragment has a m/z of 215, this fragmentation process involves also a mechanism of hydrogen transfer (Goad and Akihisa, 1997; Partridge et al., 1977; Partridge and Djerassi, 1977). This mechanism is common in all the sterols and stanols, but it is particularly present in 5α-stanols which give a strong signal for the m/z 215 ion, whereas 5β-stanols spectra have a more intense line that corresponds to the m/z $[M-90]^+$ ion and a proportionally lower line for the m/z 215 ion (Brooks et al., 1968; Goad and Akihisa, 1997).

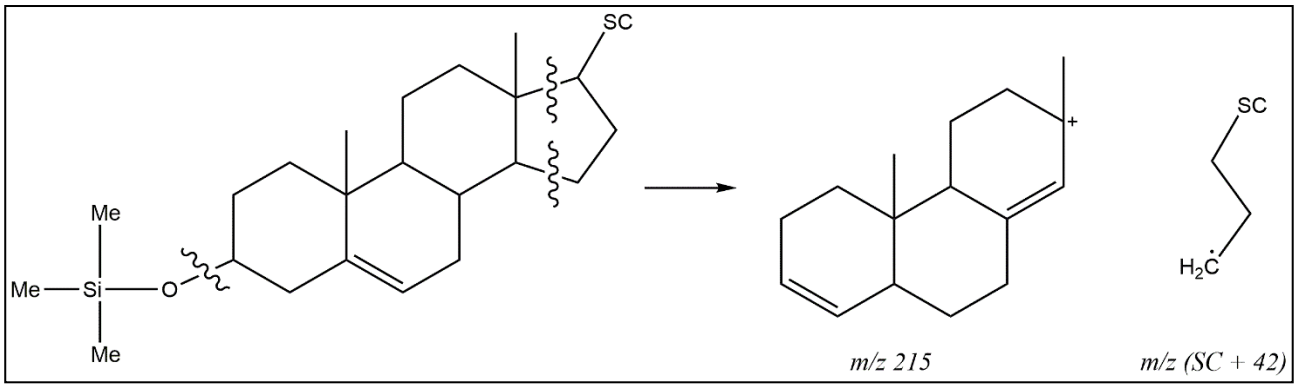

Figure 1. Breaking of the side chain and part of the D ring. Formation of the ion m/z 215.

The fragmentation across the A and B rings in the Δ^5-sterols with saturated side chain involves the breaking of the bonds: C1 – C10, C5 – C10, and C5 – C6 or C7 – C8. In the first path (with the breaking of the C5 – C6 bond) there is a loss of an ion of 157 a.m.u. resulting in $[M-157]^+$, the second fragmentation mechanism, with the cleavage of the C7 – C8 bond, brings to the molecular ions $[M-183]^+$. The resulting ions are schematically reported in figure A.4. Another mechanism involving the same part of the molecule of the Δ^5-sterols concerns the division of the C7 – C8 and C9 – C10 bonds and the loss of the side chain, producing a $[M-211]^+$ m/z ion. In addition, the rearrangement of the structure during this mechanism produces the common m/z 255 ion (Goad and Akihisa, 1997; Partridge and Djerassi, 1977; Wyllie et al., 1977).

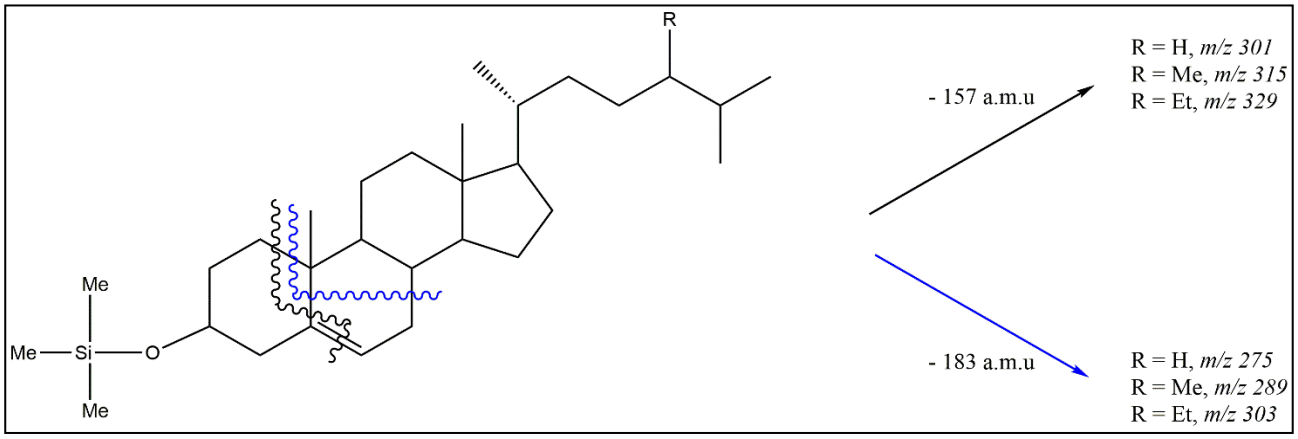

Figure A.4. Cleavage of the A and B rings of Δ^5-sterols.

Fragment of m/z 145 (or 144 if hydrogen transfer occurs) is referred to the fragment arising from the breaking of the A ring bonds C1 – C10 and C4 – C5 (Brooks et al., 1968).

Mass spectra of 5α-stanols and 5β-stanols generally show the lines corresponding to m/z 257 for the loss of the side chain and the trimehylsilanol, and m/z 230 for the loss of the side chain and part of the D ring (C16 – C17).

Appendix B

The following table B.1 reports the detailed description of the thin sections treated in the section 3.2.

Table B.1. Thin section description (terminology: Stoops 2003): [1]Thin section; [2]Stratigraphic Unit; [3]Pores: ch = channels, vu = vughs, PV = planar voids; Pa = packing voids; [4]Microstructure: ch = channel microstructure, ab = angular blocky microstructure, cm = coarse monic basic microstructure type, pl = platy, W = weakly developed, S = strongly developed. [5]Texture: CL = clay loam; [6] Organic fraction: fc = fine (<250 µm) charcoal/charred vegetal material/organic punctuations (i.e. fine black particles), cc = coarse (>250 µm) charcoal/charred vegetal material. [7]Related distribution pattern: porph = porphyric; cm = coarse monic. [8] B-fabric: cryst = crystallitic. [8]Inorganic residues of biological origin: cr = Chrysophycean stomatocysts, ph = phytoliths, eg = earthworm granules. [9]Pedofeatures: viv = vivianite crystal intergrowths; Fe hyp = Fe hypocoatings; [10]Anthropic components: slag = metal slags and spherical hammerscales; burnt soil = fragments of hearth linings & preparations; bone = fresh/digested/weathered/burnt bone fragments, fecal material = omnivore/carnivore excrements; ceramic = brick/tile/pottery fragments. Frequency determinations: percentages by visual estimate as % of the total FOV; tr = there are traces of a given feature/component, * = the feature/component is attested but does not characterise the sample/layer/horizon; ** = the feature/component is dominant and characterises the sample/layer/horizon.

Profile	**TS[1]**	**US[2]**	**Pores[3]**				**MS[4]**	**Text[5]**	**Organic[6]**		**c/f ratio**	**Rel. Dist.[6]**	**b-fabric[7]**	**Inorg. res. of biol. origin[8]**			**Pedofeatures[9]**		**Anthropic components[10]**					
			Ch	*Vu*	*PV*	*Pa*			*fc*	*cc*				*cr*	*ph*	*eg*	*Viv*	*Fe hyp*	*slag*	*burnt soil*	*mortar*	*bone*	*fecal material*	*ceramic*
SPM II	**1**	171	*	-	-	*	*ch*	*CL*	*20-30%*	*>5%*	*40/60*	*porph*	*cryst*	*	*	-	-	-	*	*	*	*	-	*
	2	171	*	-	**	-	*ch*	*CL*	*20-30%*	*5%*	*70/30*	*porph*	*cryst*	*	*	-	*tr*	-	*	*	*	*	*	*
	3	171	*	-	*	-	*ch*	*CL*	*20-30%*	*5%*	*65/35*	*porph*	*cryst*	*	*	**	-	-	*	*	*	**	*	-
	4	171	*	-	*	-	*ch*	*CL*	*20-30%*	*5-10%*	*60/40*	*porph*	*cryst*	*	-	-	-	-	-	-	*	**	-	*
	5	171	*tr*	-	*	**	*ab*	*CL*	*20%*	*5-10%*	*70/30*	*porph*	*cryst*	*	*	*	-	*tr*	-	*	*	*	*	*

		179	*Rounded gravel with mortar and brick fragments in CL matrix*																				
SPM I	**1**	165	*	-	*	-	*ch*	*CL*	*15-20%*	*5-10%*	*65/35*	*porph*	*cryst*	-	-	*tr*	-	-	**	-	*	*	*tr*	*
	2	165	*	*	**	-	*W ch*	*CL*	*15-20%*	*>5%*	*45/55*	*porph*	*cryst*	*tr*	*tr*	**	-	-	*tr*	-	*tr*	*tr*	-	**
	3	165	*	-	-	*-	*ch/cm*	*CL*	*20%*	*5-10%*	*70/30*	*porph*	*cryst*	-	-	*			*tr*	-	*	*	*tr*	*
	4	165	*	-	*	-	*ch*	*CL*	*20%*	*5-10%*	*40/60*	*porph*	*cryst*	*tr*	*	-	-	-	*	-	*tr*	*tr*	**	*
	5	165	*	-	*	--	*ch/cm*	*CL*	*15-20%*	*>5%*	*70/30*	*porph*	*cryst*	-	-	*	-	-	*tr*	-	*	*	*tr*	*
	6	165	*tr*	*	*	-	*ch/apedal*	*CL*	*20-25%*	*5%*	*45/55*	*porph*	*cryst*	*	*tr*	*	-	-	*tr*	-	*tr*	*	**	*
		179	*Rounded gravel with mortar and brick fragments in CL matrix*																				
VP I	**1**		*	*tr*	*tr*	-	*W ch*	*CL*	*20-30%*	*20-30%*	*45/55*	*porph*	*cryst*	*	*	-	-	-	**	*	*tr*	*tr*	-	*
	2		-	-	-	**	*cm*		*40%*	*40%*	-	*cm*	-	*	*	-	-	-	**	*	-	*tr*	*tr*	-
	3		*tr*	-	-	**	*cm/ab*	*CL*	*15-30%*	*10-30%*	*50/50*	*porph/cm*	*cryst*	*	*	-	-	-	**	*	*tr*	-	*	-
			tr	-	*	-	*W pl /W ab*	*CL*	*5-10%*	*>5%*	*30/70*	*porph*	*cryst*	*tr*	**	-	-	-	-	-	*tr*	*	*	-
			Mortar floor																				
VP II	**1**	613	*	-	*	-	*ch*	*CL*	*15-30%*	*20-30%*	*50/50*	*porph*	*cryst*	*tr*	*tr*	-	-	-	**	*	*tr*	**	*tr*	*
	2	613	-	-	-	**	*cm*		*40%*	*40%*	-	*cm*	-	*	*	-	*	-	**	*	**	*tr*	*tr*	-
	3A	613	*	-	-	**	*cm*	*CL*	*30%*	*30-40%*	*40/60*	*porph*	*cryst*	*	**	-	-	-	*	*	*	-	-	*tr*
		623	*Silty clay plaster/floor/hearth preparation*																				
		623	*Mortar floor*																				
	3B	624	*Burnt soil layer*																				
		624	*Charcoal layer*																				
		625	-	-	*tr*	-	*apedal*	*SaCL*	*5%*	*>2%*	*75/25*	*porph*	*cryst*	*	-	-	-	-	-	-	*	-	-	*
CDL	**1**		**	-	*	-	*ch*	*CL*	*20-25%*	*>5%*	*35/65*	*porph*	*cryst*	*	*tr*	*	-	-	-	*	-	*tr*	-	**

	2		**	-	*	-	*S ch*	*CL*	*15-20%*	*>5%*	*35/65*	*porph*	*cryst*	*	*tr*	*	-	-	-	*	-	*tr*	*tr*	*
	3		**	-	*tr*	-	*ch*	*CL*	*10-15%*	*>5%*	*35/65*	*porph*	*cryst*	*	*	*tr*	-	-	-	-	-	*tr*	-	

Appendix C

The following table C.1 reports the results of the organic carbon, TOC and ROC, and inorganic carbon (TIC), together with the results obtained from the counts and Raman characterization of charcoal fragments. The results are reported for each DE sample analyzed.

Table C.1. Results of DE samples for the amount of TOC, ROC and TIC %. Counts of charcoal classified by the mesh size dimesion (1000, 500, 250, 125, and 63 μm). Raman parameters of the position of the bands G and D and the ratio between the intensities of the two bands (I_D/I_G) and the valley intensity (I_V/I_G).

Site	Profile	Sample	TOC (%)	ROC (%)	TIC (%)	CHAR1000 (counts g^{-1})	CHAR500 (counts g^{-1})	CHAR250 (counts g^{-1})	CHAR125 (counts g^{-1})	CHAR63 (counts g^{-1})	Band D (cm^{-1})	Band G (cm^{-1})	I_D/I_G	I_V/I_G
SPM	SPM I	2 top	1.3	0.68	3.9	0	0	0	4.0	5.7				
		2 bottom	1.3	0.82	4.5	0	0	1.9	13	9.5	1311.9	1552.7	2.0	0.92
		4 top	1.5	0.93	4.5	0	0	0	5.6	7.5	1322.5	1557.6	1.7	0.86
		4 bottom	1.3	0.77	4.3	0	0	4.0	24	30				
		6 top	1.3	0.72	4.1	0	0	1.3	21	16				
		6 bottom	1.1	0.58	3.8	0	0	3.7	5.0	11	1310.4	1560.2	1.7	0.93
	SPM II	1 top	0.8	0.42	3.8	0	0	0.58	2.8	14	1300.9	1557.5	1.5	0.83
		1 bottom	1.0	0.48	3.9	0	0	1.2	5.8	11	1313.7	1554.4	1.6	0.89
		2 top	1.1	0.53	4	0	0	0.60	0	9.8	1312.5	1548.3	1.6	1.0
		2 bottom	1.3	0.63	4.2	0	0	1.3	4.4	11	1314.9	1562.3	1.9	1.1
		3 top	1.3	0.71	4.1	0	0	3.5	16	13	1308.09	1553.7	1.6	0.90
		3 bottom	1.4	0.79	4.3	0	0	7.1	3.2	12	1309.5	1551.0	1.6	0.87
		4 top	1.8	0.91	4.3	0	0	15	23	17	1314.8	1555.2	1.4	1.0
		4 bottom	1.6	0.92	4.4	0	0	2.5	9.0	16	1328.3	1551.8	1.6	1.0
		5 top	1.6	0.91	4.3	0	0	1.3	15	25	1312.4	1551.7	2.0	0.94
		5 bottom	1.5	0.8	4.4	0	0	0.46	8.6	17	1316.2	1557.6	1.8	0.89

		6 top	1.5	0.77	4.3	0	0	1.3	16	20	1317.2	1553.1	1.8	0.95
		6 bottom	0.9	0.51	4.9	0	0	3.2	3.8	12	1317.7	1552.5	1.7	0.96
VP	**VP I**	**1 top**	1.2	0.92	3.9	0	0	3.9	9.0	12	1312.3	1543.3	2.4	1.1
		1 bottom	1.4	1.3	3.6	0	0	2.3	11	19	1319.2	1560.4	1.7	0.83
		2 top	7.2	1.5	2	2.4	6.9	28	68	30	1317.2	1544.9	2.5	0.87
		2 bottom	6.7	2.4	2.2	0.95	0.99	20	64	105	1320.1	1538.8	1.5	0.99
		3 top	2.2	2.3	2.8	2.2	3.4	38	94	110	1315.9	1548.3	1.8	1.0
		3 bottom	1.2	1.0	5	4.6	4.3	30	76	83	1320.5	1548.1	1.3	0.83
	VP II	**1 top**	1.2	0.94	3.7	0	1.3	6.2	16	22	1301.2	1551.7	2.2	0.73
		1 bottom	5.4	0.91	1.8	0	1.2	21	29	27	1316.5	1549.1	1.7	0.82
		2 top	3.3	2.5	1.9	0	3.0	14	24	25				
		2 bottom	1.1	0.49	5.5	1.8	4.1	40	64	100				
		3 top	2.0	2.0	2.3	1.0	3.5	6.8	22	27	1323.6	1540.8	2.1	1.1
		3 bottom	2.0	1.5	2.8	0	0.66	7.9	24	66	1323.6	1568.7	1.7	0.83
CDL		**1**	0.7	0.39	3.6	0	0	0	0.64	14	1320.0	1545.0	1.6	1.6
		2	0.8	0.49	3.7	0	0	0	0.56	14	1313.5	1537.5	1.9	1.9
		3	1.0	0.48	3.6	0	0	0	0	10	1313.9	1547.5	1.6	1.6

Appendix D

In this appendix the concentrations of the biomarkers are reported for each DE sample analyzed. Table D.1 reports the values of the *n*-alkanes and miliacin whereas the table D.2 accounts for the PAHs analytes. The concentrations are expressed as nanogram of the compound divided by the grams of sample extracted. The values that were below the value of the blanks are reported with the caption <lod, indicating that the value is below the procedural limit of detection.

Table D.1. Concentration of miliacin and n-alkanes (expressed in in ng g^{-1}) for each DE sample analyzed. The samples are reported divided by site and profile.

Site	Profile	Sample	miliacin	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33	C34	C35
SPM	SPM I	2 top	155	290	<lod	270	<lod	<lod	<lod	<lod	34	<lod	<lod	<lod	<lod	76	<lod	<lod	320	550	310	960	890	720	350	590	210	300	26
		2 bottom	220	25	<lod	<lod	<lod	<lod	68	150	320	370	210	<lod	<lod	<lod	<lod	200.7	225	<lod	19	1010	230	850	860	630	<lod	1950	35
		4 top	230	370	<lod	160	425.8	<lod	60	<lod	<lod	<lod	<lod	<lod	34	81	22	123.6	<lod	310	140	750	260	3100	250	3500	210	1550	<lod
		4 bottom	270	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	26	<lod	<lod	<lod	<lod	<lod	<lod	<lod	120	19	300	120	1360	58	1580	21	800	<lod
		6 top	170	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	19	<lod	<lod	<lod	<lod	<lod	81	150	<lod	190	<lod	60	<lod
		6 bottom	96	250	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	130	22	640	<lod	880	<lod	440	<lod
	SPM II	1 top	210	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	120	120	<lod	95	<lod	<lod	<lod
		1 bottom	150	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	140	<lod	<lod	<lod	110	<lod	<lod	<lod	<lod	<lod	<lod
		2 top	99	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	300	<lod	<lod	140	<lod	<lod	<lod
		2 bottom	96	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	64	<lod	<lod	<lod	300	<lod	<lod	<lod	<lod	<lod	<lod
		3 top	81	21	170	<lod	<lod	<lod	330	<lod	<lod	440	<lod	<lod	<lod	<lod	44	<lod	65	680	970	<lod	850	<lod	<lod	<lod	<lod	110	720
		3 bottom	91	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	540	<lod	<lod	<lod	39	<lod	<lod	220	<lod	37	<lod
		4 top	140	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod
		4 bottom	250	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	160	100	410	190	1460	110	1560	47	810	<lod
		5 top	220	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	70	<lod	170	<lod	<lod	<lod	<lod	520	<lod	660	<lod	260	<lod
		5 bottom	310	170	<lod	<lod	<lod	<lod	160	<lod	74	<lod	23	<lod	<lod	<lod	<lod	<lod	<lod	61	<lod	270	91	1600	57	2100	37	1010	<lod

		6 top	230	170	<lod	<lod	<lod	<lod	<lod	<lod	65	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	170	220	71	695	540	1420	<lod	460	<lod	240	<lod
		6 bottom	210	<lod	<lod	<lod	<lod	<lod	<lod	<lod	135	220	300	81	40	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	340	150	660	<lod	310	<lod
VP	**VP I**	**1 top**	220	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	170	110	440	210	1600	120	1700	50	860	<lod
		1 bottom	320	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	27	140	<lod	220	<lod	43	<lod
		2 top	140	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	32	<lod	230	<lod	<lod	<lod	120	<lod	<lod	<lod	<lod	<lod	<lod
		2 bottom	120	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	31	<lod	<lod	<lod	<lod	<lod	25	220	140	390	155	880	58	1030	<lod	520	<lod
		3 top	130	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	21	77	<lod	200	<lod	41	<lod
		3 bottom	110	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	79	<lod	520	23	<lod	250	520	320	<lod	530	<lod	160	<lod
	VP II	**1 top**	170	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	72	<lod	<lod	<lod	310	<lod	120	<lod
		1 bottom	280	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	250	400	<lod	750	<lod	130	<lod
		2 top	32	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	30	<lod	72	<lod	<lod	<lod
		2 bottom	33	<lod	<lod	<lod	64	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	68	330	270	610	850	1380	190	1700	42	670	<lod
		3 top	72	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	560	<lod	<lod	<lod	23	<lod	<lod	<lod	<lod	<lod	<lod
		3 bottom	4.0	<lod	<lod	170	<lod	130	470	870	1260	1560	550	220	410	370	230	280	200	530	290	660	340	1390	180	1600	67	780	<lod
CDL		**1**	71	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	74	<lod	120	<lod	240	<lod	120	<lod	<lod
		2	170	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	430	<lod	100	<lod	<lod	27	22	140	<lod	280	<lod	180	<lod	1100
		3	230	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	210	<lod	570	<lod	390	<lod	<lod

Table D.2. Concentration (expressed in ng g^{-1}) of PAHs for each sample. The samples are reported divided by site and profile.

Site	Profile	Sample	Nap	Acnp	Acn	Flu	Phe	Ant	Fla	Py	BaA	Chr	Ret	BbF	BkF	BeP	BaP	Per	BghiP	IP	DBA
SPM	**SPM I**	**2 top**	<lod	0.92	<lod	0.17	2.2	<lod	<lod	<lod	<lod	0.0048	3.6	0.36	0.069	<lod	0.15	0.027	0.33	<lod	<lod
		2 bottom	<lod	<lod	<lod	0.18	1.7	0.019	<lod	<lod	<lod	0.33	2.5	0.41	<lod	0.077	0.37	0.067	0.57	<lod	0.0097
		4 top	7.2	0.10	0.83	1.6	91	3.0	15	17	3.6	2.0	38	26	<lod	26	9.0	<lod	40	<lod	23
		4 bottom	0.22	2.7	0.87	14	25	1.4	5.8	4.1	3.6	5.5	9.5	7.5	2.2	12	2.4	2.0	11	10	2.6
		6 top	130	11	5.4	32	6.4	0.74	3.8	3.7	2.2	4.1	7.8	4.9	1.8	9.7	1.3	2.3	7.3	7.1	0.31
		6 bottom	0.36	<lod	<lod	0.16	3.4	0.30	1.1	0.38	0.70	1.8	4.5	2.1	0.88	4.2	1.0	0.79	3.5	2.7	0.62
	SPM	**1 top**	4.0	<lod	0.31	3.7	32	1.5	8.2	11	3.6	4.6	5.6	8.5	<lod	13	2.1	<lod	7.5	11	<lod

		1 bottom	0.39	<lod	<lod	0.32	1.5	0.072	1.3	3.1	0.14	0.65	2.1	0.93	0.35	4.2	0.41	0.83	1.7	3.7	0.21
		2 top	<lod	<lod	<lod	0.50	4.8	0.14	5.3	8.1	1.1	0.80	3.7	2.0	0.85	5.1	1.8	0.38	2.8	3.0	0.028
		2 bottom	4.0	<lod	0.27	3.1	16	0.50	17	40	2.1	2.1	5.3	2.4	0.69	13	1.0	<lod	2.9	6.8	0.36
		3 top	1.2	<lod	<lod	1.8	10	<lod	1.6	0.35	0.40	1.1	6.2	0.95	0.24	0.95	0.070	1.3	<lod	<lod	<lod
		3 bottom	7.4	<lod	0.46	4.7	29	1.1	2.3	1.3	0.27	0.88	8.6	0.72	<lod	<lod	1.1	<lod	1.7	1.3	0.42
		4 top	15	0.37	8.6	4.8	14	<lod	3.0	3.6	1.2	2.9	10	3.91	<lod	<lod	4.9	<lod	5.4	<lod	<lod
		4 bottom	0.65	<lod	<lod	0.022	0.032	0.059	<lod	<lod	0.083	0.64	6.9	1.2	0.19	0.95	<lod	<lod	1.5	0.67	0.16
		5 top	0.74	<lod	0.011	1.8	20	0.37	2.2	0.58	0.36	0.88	10	1.0	0.345	1.5	0.16	<lod	1.1	0.51	0.18
		5 bottom	1.4	<lod	<lod	0.97	19	0.47	1.2	<lod	0.123	0.62	11	<lod	1.0	1.2	0.41	<lod	1.3	0.25	0.16
		6 top	4.8	<lod	<lod	1.9	11	0.49	0.88	0.35	0.23	1.1	13	1.3	0.25	0.92	0.012	2.3	0.57	<lod	0.22
		6 bottom	64	3.7	60	19	13.	0.34	1.4	1.1	1.2	0.98	11	1.5	0.42	2.5	0.11	<lod	2.0	1.9	<lod
VP	**VP I**	**1 top**	<lod	<lod	<lod	0.089	<lod	0.013	<lod	<lod	<lod	0.067	0.83	0.51	<lod	<lod	<lod	0.036	9.4	4.8	2.2
		1 bottom	<lod	<lod	<lod	0.94	3.4	0.29	0.0018	0.14	0.50	0.67	3.7	0.81	0.18	1.2	<lod	0.034	9.9	8.2	1.2
		2 top	4.2	<lod	<lod	0.71	3.8	0.082	2.3	5.2	0.31	0.87	1.8	1.4	<lod	2.6	<lod	<lod	1.9	10	0.67
		2 bottom	<lod	<lod	<lod	0.34	3.8	0.14	0.55	<lod	0.18	1.5	0.67	1.1	0.20	0.0012	<lod	<lod	0.99	0.034	0.43
		3 top	1.3	<lod	<lod	0.52	3.5	0.073	<lod	<lod	<lod	0.27	0.73	0.49	<lod	<lod	<lod	<lod	0.49	<lod	0.15
		3 bottom	<lod	<lod	<lod	0.0012	1.1	0.11	0.11	<lod	<lod	0.58	1.2	0.83	0.15	0.28	0.48	<lod	0.30	<lod	<lod
	VP II	**1 top**	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	0.0072	0.77	0.073	<lod	<lod	0.055	<lod	0.045	<lod	<lod
		1 bottom	10.1	<lod	<lod	<lod	7.5	0.62	0.23	<lod	0.055	4.0	69	2.5	0.56	<lod	1.2	<lod	0.51	<lod	<lod
		2 top	2.6	<lod	<lod	<lod	3.6	0.16	0.51	<lod	0.17	1.3	25	0.77	0.22	0.15	0.11	<lod	0.094	<lod	<lod
		2 bottom	<lod	<lod	<lod	<lod	0.097	0.038	<lod	<lod	0.0061	0.53	2.9	1.1	<lod	<lod	<lod	<lod	3.4	1.2	0.66
		3 top	9.1	0.027	0.23	2.7	18	0.61	26	61	1.8	3.4	4.4	2.4	<lod	18	0.88	<lod	6.5	35	1.0
		3 bottom	<lod	<lod	<lod	<lod	0.95	0.11	<lod	<lod	<lod	0.51	0.96	0.17	0.011	<lod	0.14	<lod	<lod	<lod	<lod
CDL		**1**	0.4	0.39	<lod	0.31	0.83	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	<lod	0.044	<lod	<lod
		2	14.3	14	1.9	2.9	6.7	2.2	0.073	1.2	2.8	0.0088	0.38	0.50	0.83	0.52	5.1	<lod	0.42	0.83	5.8
		3	6.8	6.8	<lod	<lod	4.0	<lod	0.15	1.5	0.64	0.38	1.3	1.6	3.0	1.3	2.9	1.4	1.0	2.0	1.1

Table D.1. Concentration (expressed in ng g^{-1}) of anhydrous sougars for each sample. The samples are reported divided by site and profile.

Site	Profile	Sample	LEVO	MANNO	GALACTO
SPM	SPM I	**2 top**	72	2.1	6.4
		2 bottom	68	<lod	5.1
		4 top	38	<lod	3.2
		4 bottom	17	0.1	1.3
		6 top	13	<lod	1.4
		6 bottom	21	<lod	<lod
	SPM II	**1 top**	4.9	<lod	0.3
		1 bottom	36	<lod	1.7
		2 top	3.4	<lod	<lod
		2 bottom	111	1.4	7.9
		3 top	2.5	<lod	<lod
		3 bottom	6.9	<lod	<lod
		4 top	122	2.1	14.7
		4 bottom	12	<lod	0.9
		5 top	22	<lod	<lod
		5 bottom	15	<lod	<lod
		6 top	11	<lod	2.2
		6 bottom	80	0.6	8.0
VP	VP I	**1 top**	42	2.4	4.5
		1 bottom	14	<lod	<lod
		2 top	<lod	<lod	<lod
		2 bottom	2.1	0.4	<lod
		3 top	42	2.0	6.1
		3 bottom	15	0.8	0.1
	VP II	**1 top**	17	<lod	<lod
		1 bottom	8.1	<lod	<lod
		2 top	8.9	<lod	0.2
		2 bottom	36	2.0	2.0
		3 top	18	0.8	4.8
		3 bottom	<lod	<lod	<lod
CDL		**1**	12	<lod	<lod
		2	<lod	<lod	<lod
		3	29	1.5	<lod

Appendix E

In this appendix two tables are reported. The first one, table E.1, represent the dataset of the fecal sterols and stanols of the feces form the domestica animals. The samples are grouped for animal species (donkey, horse, sheep, goat, pig, cow, and human) and as superior group, namely genera or subfamily (equine, caprine, swine, bovine, homo). Table E.2 reports the complete results of the prediction match of the classification models employed in the prediction of the mixed origin observations. The prediction match is calculated for all the possible combination of animal groups for the investigated ratios (90:10, 80:20, 70:30, 60:40). The prediction match is reported as TC (both the contributes), PC (principal contributor, i.e., the animal group present in higher percentage), and SC (secondary contributor, i.e., the animal group less present for each ratio).

Table E.1. Dataset of fecal sterols and stanols for animal species.

ANIMAL	GROUP	COP	ECOP	CHL	CHN	CHONE	24COP	24ECOP	CAMPE	STGR	β-SITO	STGN
Donkey 1	Equine	0.059	0.018	0.019	0.013	0.039	0.37	0.088	0.038	0.071	0.25	0.12
Donkey 2	Equine	0.090	0.060	0.017	0.032	0.013	0.36	0.20	0.031	0.025	0.21	0.16
Donkey 3	Equine	0.11	0.045	0.017	0.070	0.027	0.38	0.18	0.031	0.029	0.12	0.17
Donkey 4	Equine	0.061	0.015	0.024	0.0018	0.026	0.33	0.077	0.043	0.079	0.28	0.13
Donkey 5	Equine	0.087	0.030	0.0062	0.031	0.017	0.55	0.17	0.019	0.013	0.092	0.15
Donkey 6	Equine	0.062	0.017	0.015	0.037	0.030	0.41	0.085	0.021	0.038	0.19	0.18
Donkey 7	Equine	0.11	0.075	0.013	0.043	0.013	0.36	0.22	0.022	0.025	0.21	0.12
Donkey 8	Equine	0.067	0.050	0.035	0.050	0.018	0.31	0.22	0.033	0.019	0.21	0.20
Donkey 9	Equine	0.11	0.086	0.0093	0.035	0.060	0.36	0.18	0.023	0.013	0.14	0.16
Donkey 10	Equine	0.057	0.011	0.016	0.025	0.014	0.41	0.057	0.020	0.031	0.21	0.20
Donkey 11	Equine	0.063	0.015	0.0095	0.029	0.023	0.37	0.0087	0.020	0.056	0.27	0.14
Donkey 12	Equine	0.035	0.013	0.010	0.021	0.014	0.44	0.069	0.014	0.024	0.21	0.22
Donkey 13	Equine	0.056	0.044	0.011	0.033	0.020	0.49	0.20	0.027	0.022	0.13	0.16
Donkey 14	Equine	0.11	0.051	0.033	0.037	0.033	0.34	0.12	0.033	0.033	0.19	0.14
Donkey 15	Equine	0.081	0.020	0.020	0.028	0.041	0.39	0.065	0.024	0.053	0.17	0.17
Donkey 16	Equine	0.098	0.059	0.020	0.039	0.028	0.39	0.19	0.020	0.024	0.13	0.19
Donkey 17	Equine	0.10	0.045	0.025	0.037	0.033	0.36	0.15	0.041	0.045	0.14	0.17
Donkey18	Equine	0.13	0.035	0.041	0.11	0.020	0.28	0.088	0.021	0.026	0.18	0.16
Horse 1	Equine	0.069	0.022	0.025	0.037	0.036	0.40	0.25	0.032	0.025	0.17	0.18
Horse 2	Equine	0.047	0.019	0.016	0.024	0.0098	0.42	0.17	0.037	0.037	0.23	0.16
Horse 3	Equine	0.054	0.057	0.014	0.025	0.019	0.32	0.090	0.033	0.022	0.25	0.20
Horse 4	Equine	0.066	0.022	0.013	0.021	0.012	0.43	0.096	0.034	0.031	0.19	0.18
Horse 5	Equine	0.11	0.0087	0.037	0.020	0.050	0.45	0.19	0.0063	0.0042	0.055	0.26
Horse 6	Equine	0.047	0.017	0.016	0.015	0.0099	0.44	0.071	0.037	0.034	0.22	0.16
Horse 7	Equine	0.084	0.032	0.013	0.018	0.020	0.49	0.21	0.026	0.022	0.14	0.15

Horse 8	Equine	0.11	0.0024	0.027	0.012	0.033	0.47	0.26	0.0063	0.0051	0.060	0.27
Horse 9	Equine	0.051	0.057	0.0077	0.017	0.049	0.53	0.28	0.018	0.013	0.13	0.13
Horse 10	Equine	0.038	0.025	0.010	0.019	0.028	0.49	0.19	0.027	0.020	0.18	0.16
Horse 11	Equine	0.085	0.029	0.018	0.020	0.027	0.43	0.18	0.035	0.035	0.16	0.16
Horse 12	equino	0.096	0.072	0.031	0.027	0.031	0.32	0.19	0.034	0.034	0.19	0.16
Sheep 1	Caprine	0.15	0.0042	0.045	0.021	0.061	0.42	0.036	0.0084	0.0073	0.087	0.20
Sheep 2	Caprine	0.15	0.0078	0.038	0.028	0.093	0.41	0.063	0.0066	0.0046	0.062	0.19
Sheep 3	Caprine	0.075	0.012	0.013	0.017	0.042	0.32	0.15	0.0019	0.0045	0.060	0.45
Sheep 4	Caprine	0.11	0.0052	0.021	0.017	0.056	0.48	0.079	0.0081	0.0091	0.076	0.22
Sheep 5	Caprine	0.12	0.014	0.023	0.028	0.028	0.41	0.15	0.0056	0.0090	0.055	0.31
Sheep 6	Caprine	0.11	0.0058	0.024	0.049	0.066	0.41	0.063	0.0086	0.011	0.074	0.24
Sheep 7	Caprine	0.16	0.021	0.046	0.014	0.092	0.39	0.12	0.0080	0.0084	0.090	0.17
Sheep 8	Caprine	0.12	0.013	0.028	0.060	0.031	0.25	0.13	0.0076	0.0062	0.068	0.41
Sheep 9	Caprine	0.26	0.0079	0.0019	0.0016	0.090	0.42	0.045	0.0054	0.0050	0.057	0.15
Sheep 10	Caprine	0.14	0.0072	0.072	0.021	0.069	0.40	0.031	0.0095	0.014	0.071	0.19
Sheep 11	Caprine	0.097	0.011	0.043	0.024	0.075	0.32	0.10	0.0081	0.013	0.051	0.35
Sheep 12	Caprine	0.058	0.0058	0.025	0.021	0.072	0.44	0.065	0.010	0.014	0.087	0.26
Sheep 13	Caprine	0.19	0.011	0.043	0.024	0.11	0.31	0.056	0.0075	0.011	0.041	0.25
Sheep 14	Caprine	0.084	0.011	0.019	0.051	0.045	0.39	0.15	0.0053	0.0053	0.064	0.32
Sheep 15	Caprine	0.13	0.014	0.036	0.066	0.061	0.27	0.082	0.0070	0.011	0.062	0.34
Goat 1	Caprine	0.19	0.074	0.10	0.061	0.036	0.16	0.25	0.015	0.012	0.093	0.26
Goat 2	Caprine	0.21	0.034	0.0073	0.035	0.027	0.53	0.17	0.0018	0.0010	0.0086	0.15
Goat 3	Caprine	0.074	0.031	0.017	0.062	0.057	0.23	0.090	0.0032	0.0017	0.025	0.50
Goat 4	Caprine	0.28	0.030	0.030	0.013	0.15	0.39	0.091	0.00070	0.0030	0.079	0.013
Goat 5	Caprine	0.30	0.056	0.024	0.068	0.063	0.18	0.19	0.0053	0.0050	0.029	0.27
Goat 6	Caprine	0.25	0.030	0.052	0.014	0.17	0.32	0.071	0.0027	0.0037	0.039	0.11
Goat 7	Caprine	0.16	0.036	0.030	0.057	0.093	0.38	0.21	0.0052	0.0070	0.027	0.20
Goat 8	Caprine	0.26	0.017	0.060	0.025	0.10	0.28	0.26	0.0023	0.0033	0.047	0.20
Goat 9	Caprine	0.081	0.015	0.018	0.026	0.030	0.32	0.28	0.0038	0.0045	0.052	0.45
Goat 10	Caprine	0.062	0.023	0.016	0.024	0.077	0.33	0.19	0.0036	0.0045	0.055	0.40
Goat 11	Caprine	0.12	0.031	0.012	0.031	0.056	0.54	0.18	0.0047	0.0054	0.075	0.11
Goat 12	Caprine	0.21	0.018	0.069	0.026	0.053	0.15	0.19	0.0079	0.011	0.053	0.40
Goat 13	Caprine	0.17	0.042	0.035	0.034	0.14	0.37	0.25	0.0058	0.0088	0.022	0.17
Goat 14	Caprine	0.17	0.021	0.020	0.060	0.020	0.38	0.17	0.0065	0.0044	0.051	0.27
Pig 1	Swine	0.21	0.079	0.10	0.0040	0.097	0.24	0.20	0.026	0.013	0.16	0.059
Pig 2	Swine	0.29	0.011	0.072	0.011	0.092	0.24	0.12	0.029	0.013	0.17	0.071
Pig 3	Swine	0.23	0.16	0.028	0.022	0.068	0.29	0.19	0.0095	0.014	0.079	0.11
Pig 4	Swine	0.34	0.095	0.023	0.017	0.066	0.32	0.076	0.011	0.0070	0.050	0.072
Pig 5	Swine	0.29	0.13	0.023	0.011	0.069	0.34	0.019	0.0053	0.0077	0.063	0.064
Pig 6	Swine	0.29	0.13	0.053	0.017	0.16	0.14	0.027	0.019	0.025	0.14	0.021
Pig 7	Swine	0.20	0.023	0.079	0.0086	0.026	0.24	0.047	0.051	0.036	0.24	0.083
Pig 8	Swine	0.24	0.015	0.077	0.022	0.022	0.27	0.018	0.045	0.027	0.18	0.10
Pig 9	Swine	0.26	0.12	0.056	0.018	0.15	0.12	0.070	0.038	0.031	0.17	0.023
Pig 10	Swine	0.39	0.15	0.026	0.015	0.11	0.24	0.080	0.0058	0.0073	0.038	0.011
Pig 11	Swine	0.39	0.15	0.029	0.019	0.11	0.23	0.076	0.0097	0.0084	0.039	0.013
Pig 12	Swine	0.23	0.14	0.020	0.011	0.10	0.32	0.22	0.012	0.0092	0.066	0.074
Cow 1	Bovine	0.12	0.027	0.065	0.022	0.054	0.38	0.15	0.011	0.049	0.092	0.17
Cattle 2	Bovine	0.16	0.039	0.096	0.034	0.042	0.26	0.21	0.016	0.039	0.093	0.21

Cattle 3	Bovine	0.12	0.020	0.036	0.011	0.044	0.49	0.15	0.014	0.016	0.10	0.15
Cattle 4	Bovine	0.32	0.019	0.070	0.017	0.13	0.29	0.079	0.0072	0.012	0.046	0.090
Cattle 5	Bovine	0.057	0.011	0.057	0.011	0.64	0.045	0.067	0.00001	0.079	0.011	0.091
Cattle 6	Bovine	0.30	0.018	0.059	0.018	0.13	0.31	0.046	0.0042	0.013	0.056	0.097
Cattle 7	Bovine	0.38	0.014	0.080	0.017	0.13	0.24	0.28	0.0091	0.012	0.044	0.070
Cattle 8	Bovine	0.10	0.045	0.020	0.0035	0.17	0.49	0.28	0.011	0.020	0.088	0.048
Cattle 9	Bovine	0.33	0.060	0.15	0.023	0.15	0.12	0.25	0.0085	0.0071	0.048	0.10
Cattle 10	Bovine	0.25	0.030	0.18	0.039	0.10	0.12	0.27	0.012	0.012	0.061	0.19
Cattle 11	Bovine	0.20	0.048	0.22	0.053	0.065	0.13	0.13	0.011	0.014	0.067	0.19
Cattle 12	Bovine	0.29	0.019	0.023	0.0038	0.13	0.45	0.15	0.0032	0.0038	0.035	0.041
Cattle 13	Bovine	0.12	0.012	0.034	0.0081	0.056	0.54	0.381	0.0081	0.011	0.089	0.12
Cattle 14	Bovine	0.20	0.012	0.076	0.032	0.45	0.037	0.23	0.0025	0.027	0.025	0.14
Cattle 15	Bovine	0.20	0.070	0.079	0.020	0.042	0.30	0.15	0.020	0.022	0.092	0.16
Human 1	Homo	0.63	0.0079	0.069	0.0087	0.041	0.19	nd	0.0018	0.0046	0.028	0.015
Human 2	Homo	0.68	0.0077	0.083	0.0083	0.032	0.14	nd	0.0022	0.0042	0.024	0.012
Human 3	Homo	0.62	0.0068	0.071	0.0056	0.035	0.22	nd	0.0012	0.0048	0.023	0.012
Human 4	Homo	0.62	0.0094	0.053	0.013	0.049	0.23	nd	0.0022	0.0033	0.022	0.0013
Human 5	Homo	0.58	0.010	0.066	0.014	0.054	0.24	nd	0.0030	0.0038	0.030	0.0016
Human 6	Homo	0.58	0.016	0.084	0.016	0.059	0.21	nd	0.0027	0.0027	0.024	0.0016
Human 7	Homo	0.68	0.011	0.045	0.014	0.061	0.16	nd	0.0017	0.0037	0.018	0.0011
Human 8	Homo	0.63	0.0064	0.062	0.011	0.037	0.22	nd	0.0017	0.0022	0.028	0.0014
Human 9	Homo	0.58	0.0065	0.061	0.014	0.061	0.24	nd	0.0030	0.0029	0.027	0.0017
Human 10	Homo	0.63	0.0079	0.069	0.0087	0.041	0.19	nd	0.0018	0.0046	0.028	0.015

Table E.2. prediction match of the mixed contribution dataset for each model tested.

				LDA			**tree**			**RF**		
PC	**Conc PC**	**SC**	**Conc SC**	**TC (%)**	**PC (%)**	**SC (%)**	**TC (%)**	**PC (%)**	**SC (%)**	**TC (%)**	**PC (%)**	**SC (%)**
Equine	90	Caprine	10	87	97	87	0	93	0	37	100	37
Equine	90	Bovine	10	0	93	7	0	93	0	63	93	63
Equine	90	Swine	10	40	100	40	90	90	90	27	97	27
Equine	90	Homo	10	0	90	0	0	70	3	0	90	0
Caprine	90	Equine	10	53	100	53	10	100	10	57	100	57
Caprine	90	Bovine	10	47	87	47	17	87	17	53	97	53
Caprine	90	Swine	10	0	87	0	60	83	60	3	100	3
Caprine	90	Homo	10	0	97	0	0	97	0	3	100	3
Bovine	90	Equine	10	0	73	7	53	87	53	23	93	23
Bovine	90	Caprine	10	80	83	80	20	67	20	90	97	90
Bovine	90	Swine	10	10	80	10	0	93	0	30	100	30
Bovine	90	Homo	10	10	90	10	0	90	0	0	100	0
Swine	90	Equine	10	57	97	57	0	100	0	10	100	10
Swine	90	Caprine	10	3	80	3	0	87	0	17	100	17
Swine	90	Bovine	10	40	100	40	0	97	0	80	100	80
Swine	90	Homo	10	0	100	0	100	100	100	3	100	3

Homo	90	Equine	10	0	100	0	0	100	0	0	100	0
Homo	90	Caprine	10	0	100	0	0	100	0	17	100	17
Homo	90	Bovine	10	100	100	100	0	100	0	60	100	60
Homo	90	Swine	10	3	100	3	100	100	100	30	100	30
Equine	80	Caprine	20	97	100	97	0	87	0	53	100	53
Equine	80	Bovine	20	7	90	13	0	83	7	37	83	43
Equine	80	Swine	20	60	90	60	83	83	90	20	87	20
Equine	80	Homo	20	0	97	0	0	13	53	0	73	0
Caprine	80	Equine	20	77	100	77	3	83	13	37	77	37
Caprine	80	Bovine	20	60	97	60	30	97	30	77	97	77
Caprine	80	Swine	20	3	87	7	60	83	60	7	97	7
Caprine	80	Homo	20	0	90	0	0	87	0	3	97	3
Bovine	80	Equine	20	7	77	13	53	80	53	20	100	20
Bovine	80	Caprine	20	73	73	73	20	63	20	77	80	77
Bovine	80	Swine	20	10	80	10	0	83	7	40	97	40
Bovine	80	Homo	20	3	83	3	0	93	0	0	100	0
Swine	80	Equine	20	87	87	87	0	87	13	37	97	37
Swine	80	Caprine	20	7	80	7	0	90	0	27	100	27
Swine	80	Bovine	20	47	90	47	0	90	0	73	100	73
Swine	80	Homo	20	0	100	0	100	100	100	23	100	23
Homo	80	Equine	20	0	100	0	0	57	0	0	60	0
Homo	80	Caprine	20	3	100	3	0	93	7	10	100	10
Homo	80	Bovine	20	100	100	100	0	97	0	67	100	67
Homo	80	Swine	20	17	100	17	100	100	100	50	100	50
Equine	70	Caprine	30	83	87	83	0	70	0	37	73	40
Equine	70	Bovine	30	23	90	30	0	73	0	43	80	50
Equine	70	Swine	30	67	97	67	83	83	87	30	90	30
Equine	70	Homo	30	0	57	0	0	0	63	0	33	0
Caprine	70	Equine	30	73	93	73	3	57	27	30	63	30
Caprine	70	Bovine	30	70	73	70	43	87	43	70	77	70
Caprine	70	Swine	30	7	93	10	43	70	43	7	73	10
Caprine	70	Homo	30	0	93	0	0	70	0	3	97	3
Bovine	70	Equine	30	10	73	13	43	60	43	13	73	13
Bovine	70	Caprine	30	60	60	60	17	67	17	83	83	83
Bovine	70	Swine	30	27	80	33	0	83	3	50	87	50
Bovine	70	Homo	30	17	93	17	0	93	0	0	100	0
Swine	70	Equine	30	80	83	80	0	80	13	27	90	27
Swine	70	Caprine	30	20	77	20	0	57	3	23	67	23
Swine	70	Bovine	30	57	80	57	0	90	0	80	90	80
Swine	70	Homo	30	13	100	13	93	93	93	13	100	13
Homo	70	Equine	30	0	100	0	0	17	0	0	17	0
Homo	70	Caprine	30	30	100	30	0	50	27	27	60	33
Homo	70	Bovine	30	100	100	100	0	63	0	57	60	60
Homo	70	Swine	30	50	100	50	90	90	90	27	93	30

Equine	60	Caprine	40	83	83	83	0	63	0	43	77	43
Equine	60	Bovine	40	23	53	33	0	57	0	43	57	53
Equine	60	Swine	40	70	80	70	50	50	53	23	60	27
Equine	60	Homo	40	0	13	0	0	0	50	0	20	0
Caprine	60	Equine	40	67	87	67	3	73	10	20	53	23
Caprine	60	Bovine	40	53	60	53	27	60	27	57	67	57
Caprine	60	Swine	40	10	70	10	60	60	70	7	70	7
Caprine	60	Homo	40	3	77	3	0	73	0	0	87	0
Bovine	60	Equine	40	10	57	10	37	40	40	23	60	33
Bovine	60	Caprine	40	60	60	60	20	33	20	53	53	53
Bovine	60	Swine	40	30	70	30	0	80	0	67	90	67
Bovine	60	Homo	40	10	80	10	0	87	0	0	100	0
Swine	60	Equine	40	63	67	63	0	77	13	50	77	53
Swine	60	Caprine	40	23	60	33	0	50	0	17	47	20
Swine	60	Bovine	40	47	60	47	0	83	0	63	70	63
Swine	60	Homo	40	20	93	20	63	63	63	7	77	7
Homo	60	Equine	40	0	43	0	0	0	0	0	3	3
Homo	60	Caprine	40	47	83	50	0	7	47	10	17	20
Homo	60	Bovine	40	80	80	80	0	33	0	37	37	37
Homo	60	Swine	40	67	90	70	37	37	37	30	43	67

Bibliography

Alexandrovskaya, E., Alexandrovskiy, A., 2005. Radiocarbon data and anthropochemistry of ancient Moscow. Geochronometria 24, 87–95.

Alves, C.A., Gonçalves, C., Evtyugina, M., Pio, C.A., Mirante, F., Puxbaum, H., 2010. Particulate organic compounds emitted from experimental wildland fires in a Mediterranean ecosystem. Atmospheric Environment 44, 2750–2759. https://doi.org/10.1016/j.atmosenv.2010.04.029

Angelini, I., Artioli, G., Nicosia, C., 2017. Metals and metal-working residues. Archaeological Soil and Sediment Micromorphology 213–222.

Argiriadis, E., Battistel, D., McWethy, D.B., Vecchiato, M., Kirchgeorg, T., Kehrwald, N.M., Whitlock, C., Wilmshurst, J.M., Barbante, C., 2018. Lake sediment fecal and biomass burning biomarkers provide direct evidence for prehistoric human-lit fires in New Zealand. Scientific Reports 8, 2–10. https://doi.org/10.1038/s41598-018-30606-3

Argiriadis, E., Bortolini, M., Kehrwald, N.M., Roman, M., Turetta, C., Hanif, S., Erhenhi, E.O., Ramirez Aliaga, J.M., McWethy, D.B., Myrbo, A.E., Pauchard, A., Barbante, C., Battistel, D., 2021. Rapa Nui (Easter Island) Rano Raraku crater lake basin: Geochemical characterization and implications for the Ahu-Moai Period. PLoS ONE 16, 1–23. https://doi.org/10.1371/journal.pone.0254793

Arroyo-Kalin, M., 2020. Anthropogenic Sediments and Soils: Geoarchaeology. Encyclopedia of Global Archaeology 425–430. https://doi.org/10.1007/978-3-030-30018-0_856

Asare, M.O., Horák, J., Šmejda, L., Janovský, M., Hejcman, M., 2021. A medieval hillfort as an island of extraordinary fertile Archaeological Dark Earth soil in the Czech Republic. European Journal of Soil Science 72, 98–113. https://doi.org/10.1111/ejss.12965

Ascough, P.L., Bird, M.I., Scott, A.C., Collinson, M.E., Cohen-Ofri, I., Snape, C.E., Le Manquais, K., 2010. Charcoal reflectance measurements: Implications for structural characterization and assessment of diagenetic alteration. Journal of Archaeological Science 37, 1590–1599. https://doi.org/10.1016/j.jas.2010.01.020

Atkinson, C.J., Fitzgerald, J.D., Hipps, N.A., 2010. Potential mechanisms for achieving agricultural benefits from biochar application to temperate soils: A review. Plant and Soil 337, 1–18. https://doi.org/10.1007/s11104-010-0464-5

Baltensperger, D.D., 2002. Progress with proso, pearl and other millets. Trends in new crops and new uses 100–103.

Barbaro, E., Kirchgeorg, T., Zangrando, R., Vecchiato, M., Piazza, R., Barbante, C., Gambaro, A., 2015. Sugars in Antarctic aerosol. Atmospheric Environment 118, 135–144. https://doi.org/10.1016/j.atmosenv.2015.07.047

Bari, M.A., Baumbach, G., Kuch, B., Scheffknecht, G., 2009. Wood smoke as a source of particle-phase organic compounds in residential areas. Atmospheric Environment 43, 4722–4732. https://doi.org/10.1016/j.atmosenv.2008.09.006

Battistel, D., Argiriadis, E., Kehrwald, N., Spigariol, M., Russell, J.M., Barbante, C., 2017. Fire and human record at Lake Victoria, East Africa, during the Early Iron Age: Did humans or climate cause massive ecosystem changes? The Holocene 27, 997–1007. https://doi.org/10.1177/0959683616678466

Battistel, D., Kehrwald, N.M., Zennaro, P., Pellegrino, G., Barbaro, E., Zangrando, R., Pedeli, X.X., Varin, C., Spolaor, A., Vallelonga, P.T., Gambaro, A., Barbante, C., 2018. High-latitude Southern Hemisphere fire history during the mid- to late Holocene (6000-750BP). Climate of the Past 14, 871–886. https://doi.org/10.5194/cp-14-871-2018

Battistel, D., Piazza, R., Argiriadis, E., Marchiori, E., Radaelli, M., Barbante, C., 2015. GC-MS method for determining faecal sterols as biomarkers of human and pastoral animal presence in freshwater sediments. Analytical and Bioanalytical Chemistry 407, 8505–8514. https://doi.org/10.1007/s00216-015-8998-2

Beckmann, T., 1997. Preparation of thin sections for micromorphological research. Hohenheimer Bodenkundliche Hefte 40, 232.

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227. https://doi.org/10.1007/s11749-016-0481-7

Birk, J.J., Teixeira, W.G., Neves, E.G., Glaser, B., 2011. Faeces deposition on Amazonian Anthrosols as assessed from 5β-stanols. Journal of Archaeological Science 38, 1209–1220. https://doi.org/10.1016/j.jas.2010.12.015

Björkhem, I., Gustafsson, J., 1971. Mechanism of Microbial Transformation of Cholesterol into Coprostanol. European Journal of Biochemistry 21, 428–432. https://doi.org/10.1111/j.1432-1033.1971.tb01488.x

Borderie, Q., Devos, Y., Nicosia, C., Cécilia, C., Macphail, R.I., 2015. Dark Earth in the geoarchaeological approach to urban contexts. French geoarchaeology in the 21st century 213–224.

Bossard, N., Jacob, J., Le Milbeau, C., Sauze, J., Terwilliger, V., Poissonnier, B., Vergès, E., 2013. Distribution of miliacin (olean-18-en-3β-ol methyl ether) and related compounds in broomcorn millet (Panicum miliaceum) and other reputed sources: Implications for the use of sedimentary miliacin as a tracer of millet. Organic Geochemistry 63, 48–55. https://doi.org/10.1016/j.orggeochem.2013.07.012

Bourliva, A., Papadopoulou, L., Aidona, E., 2016. Study of road dust magnetic phases as the main carrier of potentially harmful trace elements. Science of the Total Environment 553, 380–391. https://doi.org/10.1016/j.scitotenv.2016.02.149

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brogiolo, G.P., Cremaschi, M., Gelichi, S., 1988. Processi di stratificazione in centri urbani (dalla stratificazione naturale alla stratificazione archeologica).

Brogiolo, G. Pietro, 1987. A proposito dell'organizzazione urbana nell'altomedioevo. A proposito dell'organizzazione urbana nell'altomedioevo 27–46.

Brönnimann, D., Röder, B., Spichtig, N., Rissanen, H., Lassau, G., Rentzel, P., 2020. The Hidden Midden: Geoarchaeological investigation of sedimentation processes, waste disposal practices, and resource management at the La Tène settlement of Basel-Gasfabrik (Switzerland). Geoarchaeology 35, 522–544. https://doi.org/10.1002/gea.21787

Brooks, C.J.W., Horning, E.C., Young, J.S., 1968. Characterization of sterols by gas chromatography-mass spectrometry of the trimethylsilyl ethers. Lipids 3, 391–402. https://doi.org/10.1007/BF02531277

Bruno, B., Fresco, P., 2014. Verona: archeologia urbana nei negozi del centro storico: lo scavo presso il nuovo store Benetton. Verona: archeologia urbana nei negozi del centro storico: lo scavo presso il nuovo store Benetton 103–111.

Bucheli, T.D., Blum, F., Desaules, A., Gustafsson, Ö., 2004. Polycyclic aromatic hydrocarbons, black carbon, and molecular markers in soils of Switzerland. Chemosphere 56, 1061–1076. https://doi.org/10.1016/j.chemosphere.2004.06.002

Bull, I.D., Lockheart, M.J., Elhmmali, M.M., Roberts, D.J., Evershed, R.P., 2002. The origin of faeces by means of biomarker detection. Environment International 27, 647–654. https://doi.org/10.1016/S0160-4120(01)00124-6

Bullock, P., Fedoroff, N., Jongerius, A., Stoops, G., Tursina, T., 1985. Handbook for soil thin section description. Waine Research.

Callegaro, A., Battistel, D., Kehrwald, N.M., Matsubara Pereira, F., Kirchgeorg, T., Del Carmen Villoslada Hidalgo, M., Bird, B.W., Barbante, C., 2018. Fire, vegetation, and Holocene climate in a southeastern Tibetan lake: A multi-biomarker reconstruction from Paru Co. Climate of the Past 14, 1543–1563. https://doi.org/10.5194/cp-14-1543-2018

Cammas, C., 2004. Les "terres noires" urbaines du Nord de la France: première typologie pédosédimentaire. Terres Noires–Dark Earth. Actes de la table ronde internationale tenue à Louvain-la-Neuve les 9, 43–55.

Cançado, L.G., Takai, K., Enoki, T., Endo, M., Kim, Y.A., Mizusaki, H., Jorio, A., Coelho, L.N., Magalhães-Paniago, R., Pimenta, M.A., 2006. General equation for the determination of the crystallite size la of nanographite by Raman spectroscopy. Applied Physics Letters 88, 1–4. https://doi.org/10.1063/1.2196057

Canuel, E.A., Martens, C.S., 1993. Seasonal variations in the sources and alteration of organic matter associated with recently-deposited sediments. Organic Geochemistry 20, 563–577. https://doi.org/10.1016/0146-6380(93)90024-6

Carreira, R.S., Wagener, A.L.R., Readman, J.W., 2004. Sterols as markers of sewage contamination in a tropical urban estuary (Guanabara Bay, Brazil): Space-time variations. Estuarine, Coastal and Shelf Science 60, 587–598. https://doi.org/10.1016/j.ecss.2004.02.014

Carver, M.O.H., 1987. The nature of urban deposits. Urban archaeology in Britain 9–26.

Caseiro, A., Bauer, H., Schmidl, C., Pio, C.A., Puxbaum, H., 2009. Wood burning impact on PM10 in three Austrian regions. Atmospheric Environment 43, 2186–2195. https://doi.org/10.1016/j.atmosenv.2009.01.012

Castelletti, L., Castiglioni, E., Rottoli, M., 2001. L'agricoltura dell'Italia settentional dal Neolitico al Medioevo. Le piante coltivate e la loro storia. FrancoAngeli, Torino 33–84.

Chan, K.H., Lam, M.H.W., Poon, K.F., Yeung, H.Y., Chiu, T.K.T., 1998. Application of sedimentary fecal stanols and sterols in tracing sewage pollution in coastal waters. Water Research 32, 225–235. https://doi.org/10.1016/S0043-1354(97)00175-9

Chandra, S., Sharma, R., Singh, K., Sharma, A., 2013. Application of bioremediation technology in the environment contaminated with petroleum hydrocarbon. Annals of microbiology 63, 417–431.

Cheng, C.H., Lehmann, J., Engelhard, M.H., 2008. Natural oxidation of black carbon in soils: Changes in molecular form and surface charge along a climosequence. Geochimica et Cosmochimica Acta 72, 1598–1610. https://doi.org/10.1016/j.gca.2008.01.010

Chia, C.H., Gong, B., Joseph, S.D., Marjo, C.E., Munroe, P., Rich, A.M., 2012. Imaging of mineral-enriched biochar by FTIR, Raman and SEM-EDX. Vibrational Spectroscopy 62, 248–257. https://doi.org/10.1016/j.vibspec.2012.06.006

Cockerton, H.E., Street-Perrott, F.A., Barker, P.A., Leng, M.J., Sloane, H.J., Ficken, K.J., 2015. Orbital forcing of glacial/interglacial variations in chemical weathering and silicon cycling within the upper White Nile basin, East Africa: Stable-isotope and biomarker evidence from Lakes Victoria and Edward. Quaternary Science Reviews 130, 57–71. https://doi.org/10.1016/j.quascirev.2015.07.028

Cohen-Ofri, I., Weiner, L., Boaretto, E., Mintz, G., Weiner, S., 2006. Modern and fossil charcoal: Aspects of structure and diagenesis. Journal of Archaeological Science 33, 428–439. https://doi.org/10.1016/j.jas.2005.08.008

Conedera, M., Tinner, W., Neff, C., Meurer, M., Dickens, A.F., Krebs, P., 2009a. Reconstructing past fire regimes: methods, applications, and relevance to fire management and conservation. Quaternary Science Reviews 28, 555–576. https://doi.org/10.1016/j.quascirev.2008.11.005

Conedera, M., Tinner, W., Neff, C., Meurer, M., Dickens, A.F., Krebs, P., 2009b. Reconstructing past fire regimes: methods, applications, and relevance to fire management and conservation. Quaternary Science Reviews 28, 555–576. https://doi.org/10.1016/j.quascirev.2008.11.005

Crabtree, P.J., Reilly, E., Wouters, B., Devos, Y., Bellens, T., Schryvers, A., 2017. Environmental evidence from early urban Antwerp: New data from archaeology, micromorphology, macrofauna and insect remains. Quaternary International 460, 108–123. https://doi.org/10.1016/j.quaint.2017.08.059

Cranwell, P.A., Eglinton, G., Robinson, N., 1987. Lipids of aquatic organisms as potential contributors to lacustrine sediments-II. Organic Geochemistry 11, 513–527. https://doi.org/10.1016/0146-6380(87)90007-6

Cremaschi, M., Nicosia, C., 2010. Corso Porta Reno, Ferrara (Northern Italy): A study in the formation processes of urban deposits. Alpine and Mediterranean Quaternary 23, 373–385.

Cuesta, A., Dhamelincourt, P., Laureyns, J., Martínez-Alonso, A., Tascón, J.M.D., 1998. Comparative performance of X-ray diffraction and Raman microprobe techniques for the study of carbon materials. Journal of Materials Chemistry 8, 2875–2879. https://doi.org/10.1039/a805841e

D'Anjou, R.M., Bradley, R.S., Balascio, N.L., Finkelstein, D.B., 2012. Climate impacts on human settlement and agricultural activities in northern Norway revealed through sediment biogeochemistry. Proceedings of the National Academy of Sciences of the United States of America 109, 20332–20337. https://doi.org/10.1073/pnas.1212730109

Daccò, C., Girometta, C., Asemoloye, M.D., Carpani, G., Picco, A.M., Tosi, S., 2020. Key fungal degradation patterns, enzymes and their applications for the removal of aliphatic hydrocarbons in polluted soils: A review. International Biodeterioration & Biodegradation 147, 104866. https://doi.org/https://doi.org/10.1016/j.ibiod.2019.104866

Dambrine, E., Dupouey, J.-L., Laüt, L., Humbert, L., Thinon, M., Beaufils, T., Richard, H., 2007. Present forest biodiersity patterns in France related to former roman agriculture. Ecology 88, 1430–1439. https://doi.org/https://doi.org/10.1890/05-1314

David, C., 2004. Les «terres noires»: outils méthodologiques, propositions analytiques et perspectives à partir de quelques exemples de sites à «terres noires» de l'Antiquité tardive et du haut Moyen-Age du nord de la France. Verslype L., Brulet R., Dark earth-Terres noires, Actes de la table ronde de Louvain-la-Neuve, Louvainla-Neuve, Université catholique de Louvain 12–31.

Davidson, D.A., Dercon, G., Stewart, M., Watson, F., 2006. The legacy of past urban waste disposal on local soils. Journal of Archaeological Science 33, 778–783. https://doi.org/10.1016/j.jas.2005.10.017

de Abreu-Mota, M.A., de Moura Barboza, C.A., Bícego, M.C., Martins, C.C., 2014. Sedimentary biomarkers along a contamination gradient in a human-impacted sub-estuary in Southern Brazil: A multi-parameter approach based on spatial and seasonal variability. Chemosphere 103, 156–163. https://doi.org/10.1016/j.chemosphere.2013.11.052

De Sousa, D.V., Guimarães, L.M., Félix, J.F., Ker, J.C., Schaefer, C.E.R.G., Rodet, M.J., 2020. Dynamic of the structural alteration

of biochar in ancient Anthrosol over a long timescale by Raman spectroscopy. PLoS ONE 15, 1–19. https://doi.org/10.1371/journal.pone.0229447

De Zanche, V., Sorbini, L., Spagna, V., 1977. Geologia del territorio del commune di Verona. Museo Civico di Storia naturale di Verona.

Deforce, K., 2017. Wood use in a growing medieval city. The overexploitation of woody resources in Ghent (Belgium) between the 10th and 12th century AD. Quaternary International 458, 123–133. https://doi.org/10.1016/j.quaint.2016.09.059

Dennison, J.R., Holtz, M., 1996. Raman spectroscopy of carbon materials. Spectroscopy (Santa Monica) 11, 38–46.

Derrien, M., Cabrera, F.A., Tavera, N.L.V., Kantún Manzano, C.A., Vizcaino, S.C., 2015. Sources and distribution of organic matter along the Ring of Cenotes, Yucatan, Mexico: Sterol markers and statistical approaches. Science of the Total Environment 511, 223–229. https://doi.org/10.1016/j.scitotenv.2014.12.053

Derrien, M., Jardé, E., Gruau, G., Pierson-Wickmann, A.C., 2011. Extreme variability of steroid profiles in cow feces and pig slurries at the regional scale: Implications for the use of steroids to specify fecal pollution sources in waters. Journal of Agricultural and Food Chemistry 59, 7294–7302. https://doi.org/10.1021/jf201040v

Derrien, M., Jardé, E., Gruau, G., Pourcher, A.M., Gourmelon, M., Jadas-Hécart, A., Pierson Wickmann, A.C., 2012. Origin of fecal contamination in waters from contrasted areas: Stanols as Microbial Source Tracking markers. Water Research 46, 4009–4016. https://doi.org/10.1016/j.watres.2012.05.003

Derrien, M., Yang, L., Hur, J., 2017. Lipid biomarkers and spectroscopic indices for identifying organic matter sources in aquatic environments: A review. Water Research 112, 58–71. https://doi.org/10.1016/j.watres.2017.01.023

Devos, Y., 2018. Near total and inorganic phosphorus concentrations as a proxy for identifying ancient activities in urban contexts: The example of dark earth in Brussels, Belgium. Geoarchaeology 33, 470–485. https://doi.org/10.1002/gea.21665

Devos, Yannick, Nicosia, C., Vrydaghs, L., Modrie, S., 2013a. Studying urban stratigraphy: Dark Earth and a microstratified sequence on the site of the Court of Hoogstraeten (Brussels, Belgium). Integrating archaeopedology and phytolith analysis. Quaternary International 315, 147–166. https://doi.org/10.1016/j.quaint.2013.07.024

Devos, Y., Nicosia, C., Vrydaghs, L., Speleers, L., van der Valk, J., Marinova, E., Claes, B., Albert, R.M., Esteban, I., Ball, T.B., Court-Picon, M., Degraeve, A., 2017. An integrated study of Dark Earth from the alluvial valley of the Senne river (Brussels, Belgium). Quaternary International 460, 175–197. https://doi.org/10.1016/j.quaint.2016.06.025

Devos, Y., Nicosia, C., Wouters, B., 2020. Urban geoarchaeology in Belgium: Experiences and innovations. Geoarchaeology 35, 27–41. https://doi.org/10.1002/gea.21755

Devos, Y., Vrydaghs, L., Degraeve, A., Fechner, K., 2009. An archaeopedological and phytolitarian study of the “Dark Earth” on the site of Rue de Dinant (Brussels, Belgium). Catena 78, 270–284. https://doi.org/10.1016/j.catena.2009.02.013

Devos, Yannick, Wouters, B., Vrydaghs, L., Tys, D., Bellens, T., Schryvers, A., 2013b. A soil micromorphological study on the origins of the early medieval trading centre of Antwerp (Belgium). Quaternary International 315, 167–183. https://doi.org/10.1016/j.quaint.2013.07.014

Dewei, R., Ling, L., Schwabacher, A.W., Young, J.W., Beitz, D.C., 1996. Mechanism of cholesterol reduction to coprostanol by Eubacterium coprostanoligenes ATCC 51222. Steroids 61, 33–40. https://doi.org/10.1016/0039-128X(95)00173-N

Diekman, J., Djerassi, C., 1967. Mass Spectrometry in Structural and Stereochemical Problems. CXXV. Mass Spectrometry of Some Steroid Trimethylsilyl Ethers. Journal of Organic Chemistry 32, 1005–1012. https://doi.org/10.1021/jo01279a033

Dresselhaus, M.S., Dresselhaus, G., Saito, R., Jorio, A., 2005. Raman spectroscopy of carbon nanotubes. Physics Reports 409, 47–99. https://doi.org/10.1016/j.physrep.2004.10.006

Eglinton, G., Gonzalez, A.G., Hamilton, R.J., Raphael, R.A., 1962. Hydrocarbon constituents of the wax coatings of plant leaves: A taxonomic survey. Phytochemistry 1, 89–102. https://doi.org/10.1016/S0031-9422(00)88006-1

Elias, V.O., Simoneit, B.R.T., Cordeiro, R.C., Turcq, B., 2001. Evaluating levoglucosan as an indicator of biomass burning in Carajás, Amazônia: A comparison to the charcoal record. Geochimica et Cosmochimica Acta 65, 267–272. https://doi.org/10.1016/S0016-7037(00)00522-6

Enache, M.D., Cumming, B.F., 2006. Tracking recorded fires using charcoal morphology from the sedimentary sequence of Prosser Lake, British Columbia (Canada). Quaternary Research 65, 282–292. https://doi.org/10.1016/j.yqres.2005.09.003

Entwistle, J.A., Abrahams, P.W., Dodgshon, R.A., 2000. The Geoarchaeological Significance and Spatial Variability of a Range

of Physical and Chemical Soil Properties from a Former Habitation Site, Isle of Skye. Journal of Archaeological Science 27, 287–303. https://doi.org/https://doi.org/10.1006/jasc.1999.0453

Entwistle, J.A., Abrahams, P.W., Dodgshon, R.A., 1998. Multi-Element Analysis of Soils from Scottish Historical Sites. Interpreting Land-Use History Through the Physical and Geochemical Analysis of Soil. Journal of Archaeological Science 25, 53–68. https://doi.org/https://doi.org/10.1006/jasc.1997.0199

Entwistle, J.A., McCaffrey, K.J.W., Dodgshon, R.A., 2007. Geostatistical and multi-elemental analysis of soils to interpret land-use history in the Hebrides, Scotland. Geoarchaeology 22, 391–415. https://doi.org/https://doi.org/10.1002/gea.20158

Evershed, R.P., Bethell, P.H., Reynolds, P.J., Walsh, N.J., 1997. 5beta-Stigmastanol and Related 5beta-Stanols as Biomarkers of Manuring : Analysis of Modern Experimental Material and. Journal of Archaeological Science 24, 485–495.

Eyssen, H.J., Parmentier, G.G., Compernolle, F.C., de Pauw, G., Piessens-Denef, M., 1973. Biohydrogenation of Sterols by Eubacterium ATCC 21,408—Nova Species. European Journal of Biochemistry 36, 411–421. https://doi.org/10.1111/j.1432-1033.1973.tb02926.x

Fabbri, D., Torri, C., Simoneit, B.R.T., Marynowski, L., Rushdi, A.I., Fabiańska, M.J., 2009. Levoglucosan and other cellulose and lignin markers in emissions from burning of Miocene lignites. Atmospheric Environment 43, 2286–2295. https://doi.org/10.1016/j.atmosenv.2009.01.030

Fang, J., Wu, F., Xiong, Y., Li, F., Du, X., An, D., Wang, L., 2014. Source characterization of sedimentary organic matter using molecular and stable carbon isotopic composition of n-alkanes and fatty acids in sediment core from Lake Dianchi, China. Science of the Total Environment 473–474, 410–421. https://doi.org/10.1016/j.scitotenv.2013.10.066

Ferrari, A.C., Robertson, J., 2004. Raman spectroscopy of amorphous, nanostructured, diamond-like carbon, and nanodiamond. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 362, 2477–2512. https://doi.org/10.1098/rsta.2004.1452

Ferrari, A.C., Robertson, J., 2000. Interpretation of Raman spectra of disordered and amorphous carbon. Physical Review B - Condensed Matter and Materials Physics 61, 14095–14107. https://doi.org/10.1103/PhysRevB.61.14095

Ficken, K.J., Li, B., Swain, D.L., Eglinton, G., 2000. An n-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. Organic Geochemistry 31, 745–749. https://doi.org/10.1016/S0146-6380(00)00081-4

Fisher, G.L., Prentice, B.A., Sllberman, D., Ondov, J.M., Biermann, A.H., Ragainl, R.C., McFarl, A.R., 1978. Physical and Morphological Studies of Size-Classified Coal Fly Ash. Environmental Science and Technology 12, 447–451. https://doi.org/10.1021/es60140a008

Galinié, H., 2007. L'expression terres noires, un concept d'attente. Les petits cahiers d'Anatole 15.

Ganzarolli, G., Alexander, M., Chavarria Arnau, A., Craig, O.E., 2018. Direct evidence from lipid residue analysis for the routine consumption of millet in Early Medieval Italy. Journal of Archaeological Science 96, 124–130. https://doi.org/10.1016/j.jas.2018.06.007

Gao, X., Norwood, M., Frederick, C., McKee, A., Masiello, C.A., Louchouarn, P., 2016. Organic geochemical approaches to identifying formation processes for middens and charcoal-rich features. Organic Geochemistry 94, 1–11. https://doi.org/10.1016/j.orggeochem.2016.01.007

Glaser, B., 2007. Prehistorically modified soils of central Amazonia: A model for sustainable agriculture in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences 362, 187–196. https://doi.org/10.1098/rstb.2006.1978

Glaser, B., Birk, J.J., 2012. State of the scientific knowledge on properties and genesis of Anthropogenic Dark Earths in Central Amazonia (terra preta de índio). Geochimica et Cosmochimica Acta 82, 39–51. https://doi.org/10.1016/j.gca.2010.11.029

Goad, L.J., Akihisa, T., 1997. Analysis of sterols, Journal of Chromatography A. https://doi.org/10.1007/978-94-009-1447-6

Golyeva, Alexandra, Chichagova, O., Bondareva, J., 2016. Soil forming processes of ancient man-made soils (cultural layers) by the example of sites in humid (Dunino) and arid (Ar-Dolong) regions of Russia: A first approach. Quaternary International 418, 22–27. https://doi.org/10.1016/j.quaint.2015.11.093

Golyeva, A, Zazovskaia, E., Turova, I., 2016. Properties of ancient deeply transformed man-made soils (cultural layers) and their advances to classification by the example of Early Iron Age sites in Moscow Region. Catena 137, 605–610.

Gourmelon, M., Caprais, M.P., Mieszkin, S., Marti, R., Wéry, N., Jardé, E., Derrien, M., Jadas-Hécart, A., Communal, P.Y.,

Jaffrezic, A., Pourcher, A.M., 2010. Development of microbial and chemical MST tools to identify the origin of the faecal pollution in bathing and shellfish harvesting waters in France. Water Research 44, 4812–4824. https://doi.org/10.1016/j.watres.2010.07.061

Grimalt, J.O., Fernández, P., Bayona, J.M., Albaigés, J., 1990. Assessment of Fecal Sterols and Ketones as Indicators of Urban Sewage Inputs to Coastal Waters. Environmental Science and Technology 24, 357–363. https://doi.org/10.1021/es00073a011

Grimalt, J.O., Torras, E., Albaigés, J., 1988. Bacterial reworking of sedimentary lipids during sample storage. Organic Geochemistry 13, 741–746. https://doi.org/10.1016/0146-6380(88)90096-4

Grossman, J.M., O'Neill, B.E., Tsai, S.M., Liang, B., Neves, E., Lehmann, J., Thies, J.E., 2010. Amazonian anthrosols support similar microbial communities that differ distinctly from those extant in adjacent, unmodified soils of the same mineralogy. Microbial Ecology 60, 192–205. https://doi.org/10.1007/s00248-010-9689-3

Gundale, M.J., DeLuca, T.H., 2006. Temperature and source material influence ecological attributes of ponderosa pine and Douglas-fir charcoal. Forest Ecology and Management 231, 86–93. https://doi.org/10.1016/j.foreco.2006.05.004

Guo, Y., Hastie, T., Tibshirani, R., 2007. Regularized linear discriminant analysis and its application in microarrays. Biostatistics 8, 86–100. https://doi.org/10.1093/biostatistics/kxj035

Gustafsson, Ö., Bucheli, T.D., Kukulska, Z., Andersson, M., Largeau, C., Rouzaud, J.N., Reddy, C.M., Eglinton, T.I., 2001. Evaluation of a protocol for the quantification of black carbon in sediments. Global Biogeochemical Cycles 15, 881–890. https://doi.org/10.1029/2000GB001380

Guttmann, E.B., Simpson, I.A., Nielsen, N., Dockrill, S.J., 2008. Anthrosols in Iron Age Shetland: Implications for arable and economic activity. Geoarchaeology: An International Journal 23, 799–823.

Haritash, A.K., Kaushik, C.P., 2009. Biodegradation aspects of Polycyclic Aromatic Hydrocarbons (PAHs): A review. Journal of Hazardous Materials 169, 1–15. https://doi.org/10.1016/j.jhazmat.2009.03.137

Harrault, L., Milek, K., Jardé, E., Jeanneau, L., Derrien, M., Anderson, D.G., 2019. Faecal biomarkers can distinguish specific mammalian species in modern and past environments. PLoS ONE 14, 1–26. https://doi.org/10.1371/journal.pone.0211119

Hart, P.E., Stork, D.G., Duda, R.O., 2000. Pattern classification. Wiley Hoboken.

Hastie, T., Tibshirani, R., James, G., Witten, D., 2021. An introduction to statistical learning (2nd ed.). Springer texts 102, 618.

Hatcher, P.G., McGillivary, P.A., 1979. Sewage Contamination in the New York Bight. Coprostanol as an Indicator. Environmental Science and Technology 13, 1225–1229. https://doi.org/10.1021/es60158a015

He, Y., Zhang, G.L., 2009. Historical record of black carbon in urban soils and its environmental implications. Environmental Pollution 157, 2684–2688. https://doi.org/10.1016/j.envpol.2009.05.019

Heimdahl, J., 2005. Urbanised Nature in the Past: Site Formation and Environmental Developement in Two Swedish Towns, AD 1200-1800.

Holliday, V.T., Gartner, W.G., 2007. Methods of soil P analysis in archaeology. Journal of Archaeological Science 34, 301–333. https://doi.org/10.1016/j.jas.2006.05.004

Howard, J.L., Orlicki, K.M., 2016. Composition, micromorphology and distribution of microartifacts in anthropogenic soils, Detroit, Michigan, USA. Catena 138, 103–116. https://doi.org/10.1016/j.catena.2015.11.016

Hunt, H. V., Campana, M.G., Lawes, M.C., Park, Y.J., Bower, M.A., Howe, C.J., Jones, M.K., 2011. Genetic diversity and phylogeography of broomcorn millet (Panicum miliaceum L.) across Eurasia. Molecular Ecology 20, 4756–4771. https://doi.org/10.1111/j.1365-294X.2011.05318.x

Hunt, H. V., Vander Linden, M., Liu, X., Motuzaite-Matuzeviciute, G., Colledge, S., Jones, M.K., 2008. Millets across Eurasia: Chronology and context of early records of the genera Panicum and Setaria from archaeological sites in the Old World. Vegetation History and Archaeobotany 17, 31–34. https://doi.org/10.1007/s00334-008-0187-1

Idler, D.R., Safe, L.M., Safe, S.H., 1970. Mass spectrometric studies of methyl ether derivatives of sterols. Steroids 16, 251–262.

Inoue, J., Yoshie, A., Tanaka, T., Onji, T., Inoue, Y., 2017. Disappearance and alteration process of charcoal fragments in cumulative soils studied using Raman spectroscopy. Geoderma 285, 164–172. https://doi.org/10.1016/j.geoderma.2016.09.032

Ioffe, S., 2006. Probabilistic linear discriminant analysis. Lecture Notes in Computer Science (including subseries Lecture Notes

in Artificial Intelligence and Lecture Notes in Bioinformatics) 3954 LNCS, 531–542. https://doi.org/10.1007/11744085_41

Ishimaru, K., Hata, T., Bronsveld, P., Meier, D., Imamura, Y., 2007a. Spectroscopic analysis of carbonization behavior of wood, cellulose and lignin. Journal of Materials Science 42, 122–129. https://doi.org/10.1007/s10853-006-1042-3

Ishimaru, K., Hata, T., Bronsveld, P., Nishizawa, T., Imamura, Y., 2007b. Characterization of sp2- and sp3-bonded carbon in wood charcoal. Journal of Wood Science 53, 442–448. https://doi.org/10.1007/s10086-007-0879-7

Isobe, K.O., Tarao, M., Zakaria, M.P., Chiem, N.H., Minh, L.Y., Takada, H., 2002. Quantitative application of fecal sterols using gas chromatography - Mass spectrometry to investigate fecal pollution in tropical waters: Western Malaysia and Mekong Delta, Vietnam. Environmental Science and Technology 36, 4497–4507. https://doi.org/10.1021/es020556h

IUSS Working Group WRB., 2014. World reference base for soil resources 2014, update 2015 - International soil classification system for naming soils and creating legends for soil maps, World Soil. ed.

Jacob, J., Disnar, J.R., Arnaud, F., Chapron, E., Debret, M., Lallier-Vergès, E., Desmet, M., Revel-Rolland, M., 2008. Millet cultivation history in the French Alps as evidenced by a sedimentary molecule. Journal of Archaeological Science 35, 814–820. https://doi.org/10.1016/j.jas.2007.06.006

Jacob, J., Disnar, J.R., Arnaud, F., Gauthier, E., Billaud, Y., Chapron, E., Bardoux, G., 2009. Impacts of new agricultural practices on soil erosion during the Bronze Age in the French Prealps. Holocene 19, 241–249. https://doi.org/10.1177/0959683608100568

Jadhav, S.D., Channe, H.P., 2016. Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. International Journal of Science and Research (IJSR) 5, 1842–1845. https://doi.org/10.21275/v5i1.nov153131

Jardé, E., Gruau, G., Mansuy-Huault, L., 2007a. Detection of manure-derived organic compounds in rivers draining agricultural areas of intensive manure spreading. Applied Geochemistry 22, 1814–1824. https://doi.org/10.1016/j.apgeochem.2007.03.037

Jardé, E., Gruau, G., Mansuy-Huault, L., Peu, P., Martinez, J., 2007b. Using sterols to detect pig slurry contribution to soil organic matter. Water, Air, and Soil Pollution 178, 169–178. https://doi.org/10.1007/s11270-006-9188-9

Jawhari, T., Roid, A., Casado, J., 1995. Raman spectroscopic characterization of some commercially available carbon black materials. Carbon 33, 1561–1565. https://doi.org/10.1016/0008-6223(95)00117-V

Jenkins, B.M., Jones, A.D., Turn, S.Q., Williams, R.B., 1996. Particle concentrations, gas-particle partitioning, and species intercorrelations for polycyclic aromatic hydrocarbons (PAH) emitted during biomass burning. Atmospheric Environment 30, 3825–3835. https://doi.org/10.1016/1352-2310(96)00084-2

Jorio, A., Ribeiro-Soares, J., Cançado, L.G., Falcão, N.P.S., Dos Santos, H.F., Baptista, D.L., Martins Ferreira, E.H., Archanjo, B.S., Achete, C.A., 2012. Microscopy and spectroscopy analysis of carbon nanostructures in highly fertile Amazonian anthrosoils. Soil and Tillage Research 122, 61–66. https://doi.org/10.1016/j.still.2012.02.009

Ju, J., Zhu, L. ping, Wang, J., Xie, M., Zhen, X., Wang, Y., Peng, P., 2010. Water and sediment chemistry of Lake Pumayum Co, South Tibet, China: Implications for interpreting sediment carbonate. Journal of Paleolimnology 43, 463–474. https://doi.org/10.1007/s10933-009-9343-6

Juste, C., Gérard, P., 2021. Cholesterol-to-coprostanol conversion by the gut microbiota: What we know, suspect, and ignore. Microorganisms 9. https://doi.org/10.3390/microorganisms9091881

Karkanas, P., Goldberg, P., 2007. Micromorphology of sediments: Deciphering archaeological context. Israel Journal of Earth Sciences 56, 63–71. https://doi.org/10.1560/IJES.56.2-4.63

Kawakami, M., Karato, T., Takenaka, T., Yokoyama, S., 2005. Structure analysis of coke, wood charcoal and bamboo charcoal by Raman spectroscopy and their reaction rate with CO2. ISIJ International 45, 1027–1034. https://doi.org/10.2355/isijinternational.45.1027

Keiluweit, M., Nico, P.S., Johnson, M., Kleber, M., 2010. Dynamic molecular structure of plant biomass-derived black carbon (biochar). Environmental Science and Technology 44, 1247–1253. https://doi.org/10.1021/es9031419

Killops, S., Killops, V., 2005. Introduction to Organic Geochemistry, 2nd edn. Blackwell Publishing Ltd. https://doi.org/10.1111/j.1468-8123.2005.00113.x

King, S.M., 2008. The spatial organization of food sharing in Early Postclassic households: an application of soil chemistry in Ancient Oaxaca, Mexico. Journal of Archaeological Science 35, 1224–1239.

Kirchgeorg, T., Schüpbach, S., Kehrwald, N., McWethy, D.B., Barbante, C., 2014. Method for the determination of specific molecular markers of biomass burning in lake sediments. Organic Geochemistry 71, 1–6. https://doi.org/10.1016/j.orggeochem.2014.02.014

Kleber, M., 2010. What is recalcitrant soil organic matter? Environmental Chemistry 7, 320–332. https://doi.org/10.1071/EN10006

Knight, D.S., White, W.B., 1989. Characterization of diamond films by Raman spectroscopy. Journal of Materials Research 4, 385–393. https://doi.org/DOI: 10.1557/JMR.1989.0385

Krupski, M., Kabala, C., Sady, A., Gliński, R., Wojcieszak, J., 2017. Double-and triple-depth digging and Anthrosol formation in a medieval and modern-era city (Wrocław, SW Poland). Geoarchaeological research on past horticultural practices. Catena 153, 9–20. https://doi.org/10.1016/j.catena.2017.01.028

Kuo, L.J., Louchouarn, P., Herbert, B.E., 2011. Influence of combustion conditions on yields of solvent-extractable anhydrosugars and lignin phenols in chars: Implications for characterizations of biomass combustion residues. Chemosphere 85, 797–805. https://doi.org/10.1016/j.chemosphere.2011.06.074

Lambrecht, G., Rodríguez de Vera, C., Jambrina-Enríquez, M., Crevecoeur, I., Gonzalez-Urquijo, J., Lazuen, T., Monnier, G., Pajović, G., Tostevin, G., Mallol, C., 2021. Characterisation of charred organic matter in micromorphological thin sections by means of Raman spectroscopy. Archaeological and Anthropological Sciences 13. https://doi.org/10.1007/s12520-020-01263-3

Leeming, R., Ball, A., Ashbolt, N., Nichols, P., 1996. Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. Water Research 30, 2893–2900. https://doi.org/10.1016/S0043-1354(96)00011-5

Leeming, R., Stark, J.S., Smith, J.J., 2014. Novel use of faecal sterols to assess human faecal contamination in Antarctica: A likelihood assessment matrix for environmental monitoring. Antarctic Science 27, 31–43. https://doi.org/10.1017/S0954102014000273

Leeming, R.L., Ball, A., Ashbolt, N.J., Jones, G., Nichols, P.D., 1994. Distinguishing between human and animal sources of faecal pollution.

Lehmann, A., Stahr, K., 2007. Nature and significance of anthropogenic urban soils. Journal of Soils and Sediments 7, 247–260. https://doi.org/10.1065/jss2007.06.235

Leogrande, R., Vitti, C., Castellini, M., Mastrangelo, M., Pedrero, F., Vivaldi, G.A., Stellacci, A.M., 2021. Comparison of two methods for total inorganic carbon estimation in three soil types in mediterranean area. Land 10, 1–11. https://doi.org/10.3390/land10040409

Leonardi, G., Miglavacca, M., Nardi, S., 1999. Soil Phosphorus Analysis as an Integrative Tool for Recognizing Buried Ancient Ploughsoils. Journal of Archaeological Science 26, 343–352. https://doi.org/https://doi.org/10.1006/jasc.1998.0329

Lima, A.L.C., Farrington, J.W., Reddy, C.M., 2005. Combustion-derived polycyclic aromatic hydrocarbons in the environment - A review. Environmental Forensics 6, 109–131. https://doi.org/10.1080/15275920590952739

Linderholm, J., 2003. Soil chemical surveying: a path to a different understanding of prehistoric sites and societies in northern Sweden, in: Second International Conference on Soils and Archaeology. Felici Editore Pisa, pp. 114–117.

Llorente, M., Glaser, B., Turrión, M.B., 2010. Storage of organic carbon and Black carbon in density fractions of calcareous soils under different land uses. Geoderma 159, 31–38. https://doi.org/10.1016/j.geoderma.2010.06.011

Loh, W.Y., 2014. Fifty years of classification and regression trees. International Statistical Review 82, 329–348. https://doi.org/10.1111/insr.12016

Lombardo, U., Arroyo-Kalin, M., Schmidt, M., Huisman, H., Lima, H.P., de Paula Moraes, C., Neves, E.G., Clement, C.R., Aires da Fonseca, J., de Almeida, F.O., Vieira Alho, C.F.B., Bronk Ramsey, C., Brown, G.G., Cavallini, M.S., Lima da Costa, M., Cunha, L., dos Anjos, L.H.C., Denevan, W.M., Fausto, C., Fernandes Caromano, C., Fontana, A., Franchetto, B., Glaser, B., Heckenberger, M.J., Hecht, S., Honorato, V., Jarosch, K.A., Braga Junqueira, A., Kater, T., Tamanaha, E.K., Kuyper, T.W., Lehmann, J., Madella, M., Maezumi, S.Y., Matthews Cascon, L., Mayle, F.E., McKey, D., Moraes, B., Morcote-Ríos, G., Palheta Barbosa, C.A., Magalhães, M.P., Prestes-Carneiro, G., Pugliese, F., Pupim, F.N., Raczka, M.F., Py-Daniel, A.R., Riris, P., Cigaran da Rocha, B., Rodrigues, L., Rostain, S., Macedo, R.S., Shock, M.P., Sprafke, T., Stampanoni Bassi, F., Valle, R., Vidal-Torrado, P., Villagrán, X.S., Watling, J., Weber, S.L., Teixeira, W.G., 2022. Evidence confirms an anthropic origin of Amazonian Dark Earths. Nature Communications 13, 1–6. https://doi.org/10.1038/s41467-022-31064-2

Louchouarn, P., Kuo, L.J., Wade, T.L., Schantz, M., 2009. Determination of levoglucosan and its isomers in size fractions of aerosol standard reference materials. Atmospheric Environment 43, 5630–5636. https://doi.org/10.1016/j.atmosenv.2009.07.040

Lu, H., Zhang, J., Liu, K.B., Wu, N., Li, Y., Zhou, K., Ye, M., Zhang, T., Zhang, H., Yang, X., Shen, L., Xu, D., Li, Q., 2009. Earliest domestication of common millet (Panicum miliaceum) in East Asia extended to 10,000 years ago. Proceedings of the National Academy of Sciences of the United States of America 106, 7367–7372. https://doi.org/10.1073/pnas.0900158106

Lu, T., Wang, X., Zhang, W., 2020. Total and dissolved soil organic and inorganic carbon and their relationships in typical loess cropland of Fengu Basin. Geoscience Letters 7. https://doi.org/10.1186/s40562-020-00167-3

Lucheta, A.R., de Souza Cannavan, F., Roesch, L.F.W., Tsai, S.M., Kuramae, E.E., 2016. Fungal Community Assembly in the Amazonian Dark Earth. Microbial Ecology 71, 962–973. https://doi.org/10.1007/s00248-015-0703-7

Macphail, R., Linderholm, J., 2004. " Dark earth": recent studies of" Dark earth" and" dark-earth-like" microstratigraphy in England, UK, in: Verslype, L. & Brulet, R. (Eds.), Terres Noires – Dark Earth, Actes de La Table-Ronde Internationale Tenue à Louvain-La-Neuve, Les 09 et 10 Novembre 2001. UCL, Louvain-la-Neuve, pp. 35–42.

Macphail, R.I., 2010. Dark earth and insights into changing land use of urban areas. Leicester Archaeology.

Macphail, R.I., 2003. Soil microstratigraphy: a micromorphological and chemical approach. MoLAS.

Macphail, R.I., 1994. The reworking of urban stratigraphy by human and natural processes. Urban-rural connexions: perspectives from environmental archaeology 13–43.

Macphail, R.I., 1981. Soil and botanical studies of the "Dark Earth," in: Jones, M.., Dimbleby, G.W. (Eds.), The Environment of Man: The Iron Age to the Anglo-Saxon Period. Oxford, pp. 309–331.

Macphail, R.I., Courty, M.A., Gebhardt, A., 1990a. Soil micromorphological evidence of early agriculture in north-west Europe. World Archaeology 22, 53–69. https://doi.org/10.1080/00438243.1990.9980129

Macphail, R.I., Courty, M.A., Goldberg, P., 1990b. Soil micromorphology in archaeology. Endeavour 14, 163–171. https://doi.org/10.1016/0160-9327(90)90039-T

Macphail, R.I., Galinié, H., Verhaeghe, F., 2003a. A future for dark earth? Antiquity 77, 349–358. https://doi.org/10.1017/S0003598X00092334

Macphail, R.I., Galinié, H., Verhaeghe, F., 2003b. A future for Dark Earth? Antiquity 77.296, 349–358.

Madgwick, R., Mulville, J., Stevens, R.E., 2012. Diversity in foddering strategy and herd management in late Bronze Age Britain: An isotopic investigation of pigs and other fauna from two midden sites. Environmental Archaeology 17, 126–140. https://doi.org/10.1179/1461410312Z.00000000011

Martins, C.C., Seyffert, B.H., Braun, J.A.F., Fillmann, G., 2011. Input of organic matter in a large South American tropical estuary (paranaguá estuarine system, Brazil) indicated by sedimentary sterols and multivariate statistical approach. Journal of the Brazilian Chemical Society 22, 1585–1594. https://doi.org/10.1590/S0103-50532011000800023

Martins, C.D.C., Fillmann, G., Montone, R.C., 2007. Natural and anthropogenic sterols inputs in surface sediments of Patos Lagoon, Brazil. Journal of the Brazilian Chemical Society 18, 106–115. https://doi.org/10.1590/S0103-50532007000100012

Mauquoy, D., Payne, R.J., Babeshko, K. V., Bartlett, R., Boomer, I., Bowey, H., Evans, C.D., Ring-Hrubesh, F., Muirhead, D., O'Callaghan, M., Piotrowska, N., Rush, G., Sloan, T., Smeaton, C., Tsyganov, A.N., Mazei, Y.A., 2020. Falkland Island peatland development processes and the pervasive presence of fire. Quaternary Science Reviews 240, 106391. https://doi.org/10.1016/j.quascirev.2020.106391

Mazurek, R., Kowalska, J., Gasiorek, M., Setlak, M., 2016. Micromorphological and physico-chemical analyses of cultural layers in the urban soil of a medieval city - A case study from Krakow, Poland. Catena 141, 73–84. https://doi.org/10.1016/j.catena.2016.02.026

McDonald-Wharry, J., Manley-Harris, M., Pickering, K., 2013. Carbonisation of biomass-derived chars and the thermal reduction of a graphene oxide sample studied using Raman spectroscopy. Carbon 59, 383–405. https://doi.org/10.1016/j.carbon.2013.03.033

McGowan, G., Prangnell, J., 2006. The significance of vivianite in archaeological settings. Geoarchaeology 21, 93–111. https://doi.org/10.1002/gea.20090

McGrath, T.E., Chan, W.G., Hajajigol, R., 2003. Low temperature mechanism for the formation of polycyclic aromatic hydrocarbons from the pyrolysis of cellulose. Journal of Analytical and Applied Pyrolysis, v. 66, p. 51-70, 2003. Journal of Analytical and Applied Pyrolysis 66, 51–70.

McLachlan, G.J., 2005. Discriminant analysis and statistical pattern recognition. John Wiley & Sons.

McWethy, D.B., Alt, M., Argiriadis, E., Battistel, D., Everett, R., Pederson, G.T., 2020. Millennial-Scale Climate and Human Drivers of Environmental Change and Fire Activity in a Dry, Mixed-Conifer Forest of Northwestern Montana. Frontiers in Forests and Global Change 3, 1–16. https://doi.org/10.3389/ffgc.2020.00044

McWethy, D.B., Whitlock, C., Wilmshurst, J.M., McGlone, M.S., Fromont, M., Li, X., Dieffenbacher-Krall, A., Hobbs, W.O., Fritz, S.C., Cook, E.R., 2010. Rapid landscape transformation in South Island, New Zealand, following initial Polynesian settlement. Proceedings of the National Academy of Sciences of the United States of America 107, 21343–21348. https://doi.org/10.1073/pnas.1011801107

Mernagh, T.P., Cooney, R.P., Johnson, R.A., 1984. Raman spectra of Graphon carbon black. Carbon 22, 39–42. https://doi.org/https://doi.org/10.1016/0008-6223(84)90130-1

Meuser, H., 2010. Anthropogenic Soils. https://doi.org/10.1007/978-90-481-9328-8_5

Meyers, P.A., 2003. Application of organic geochemistry to paleolimnological reconstruction: a summary of examples from the Laurention Great Lakes. Organic Geochemistry 34, 261–289.

Migliavacca, M., Pizzeghello, D., Ertani, A., Nardi, S., 2013. Chemical analyses of archaeological sediments identified the ancient activity areas of an Iron age building at Rotzo (Vicenza, Italy). Quaternary International 289, 101–112. https://doi.org/10.1016/j.quaint.2012.07.016

Miller, D.R., Castañeda, I.S., Bradley, R.S., MacDonald, D., 2017. Local and regional wildfire activity in central Maine (USA) during the past 900 years. Journal of Paleolimnology 58, 455–466. https://doi.org/10.1007/s10933-017-0002-z

Motuzaite-Matuzeviciute, G., Staff, R.A., Hunt, H. V., Liu, X., Jones, M.K., 2013. The early chronology of broomcorn millet (Panicum miliaceum) in Europe. Antiquity 87, 1073–1085. https://doi.org/10.1017/S0003598X00049875

Moucawi, J., Fustec, E., Jambu, P., Amblfs, A., Jacquesy, R., 1981. Biooxidation of added and natural hydrocarbons in soils: Effect of iron. Soil Biology and Biochemistry 13, 335–342. https://doi.org/10.1016/0038-0717(81)90073-0

Mudge, S.M., Duce, C.E., 2005. Identifying the source, transport path and sinks of sewage derived organic matter. Environmental Pollution 136, 209–220. https://doi.org/10.1016/j.envpol.2005.01.015

Mudge, S.M., Seguel, C.G., 1999. Organic contamination of San Vicente Bay, Chile. Marine Pollution Bulletin 38, 1011–1021. https://doi.org/10.1016/S0025-326X(99)00132-0

Murphy, C., 2016. Finding millet in the Roman world. Archaeological and Anthropological Sciences 8, 65–78. https://doi.org/10.1007/s12520-015-0237-4

Murtaugh, J.J., Bunch, R.L., 1967. Sterols as a measure of fecal pollution. Journal (Water Pollution Control Federation) 404–409.

Musa Bandowe, B.A., Srinivasan, P., Seelge, M., Sirocko, F., Wilcke, W., 2014. A 2600-year record of past polycyclic aromatic hydrocarbons (PAHs) deposition at Holzmaar (Eifel, Germany). Palaeogeography, Palaeoclimatology, Palaeoecology 401, 111–121. https://doi.org/10.1016/j.palaeo.2014.02.021

Mustaphi, C.J.C., Pisaric, M.F.J., 2014. A classification for macroscopic charcoal morphologies found in Holocene lacustrine sediments. Progress in Physical Geography 38, 734–754. https://doi.org/10.1177/0309133314548886

Nemanich, R.J., Solin, S.A., 1978. First- and second-order Raman scattering from finite-size crystals of graphite. Physical Review B 20.

Nesbitt, B.M., Summers, G.D., 2022. Some Recent Discoveries of Millet (Panicum miliaceum L . and Setaria italica (L .) P . Beauv .) at Excavations in Turkey and Iran Author (s): Mark Nesbitt and G . D . Summers Source : Anatolian Studies , Vol . 38 (1988), pp . 85-97 Published by : Br 38, 85–97.

Ngu, L.N., Wu, H., Zhang, D.K., 2007. Characterisation of cenospheres in fly ash from Australian power stations. 6th Asia-Pacific Conference on Combustion, ASPACC 2007 31, 3437–3445.

Nicosia, C., 2018. Geoarcheologia delle stratificazioni urbane post-classiche, Antenor Qu. ed. Edizioni Quasar.

Nicosia, C., Balista, C., Cupitò, M., Ertani, A., Leonardi, G., Nardi, S., Vidale, M., 2011. Anthropogenic deposits from the Bronze Age site of Fondo Paviani (Verona, Italy): Pedochemical and micropedological characteristics. Quaternary International

243, 280–292. https://doi.org/10.1016/j.quaint.2010.10.006

Nicosia, C., Devos, Y., 2014. Urban dark earth, in: Encyclopedia of Global Archaeology. Springer, pp. 7532–7540.

Nicosia, C., Devos, Y., Macphail, R.I., 2017. European dark earth. Archaeological soil and sediment micromorphology 331–343.

Nicosia, C., Ertani, A., Vianello, A., Nardi, S., Brogiolo, G. Pietro, Chavarría Arnau, A., Becherini, F., 2019. Heart of darkness: an interdisciplinary investigation of the urban anthropic deposits of the Baptistery of Padua (Italy). Archaeological and Anthropological Sciences 11, 1977–1993. https://doi.org/10.1007/s12520-018-0646-2

Nicosia, C., Langohr, R., Mees, F., Arnoldus-Huyzendveld, A., Bruttini, J., Cantini, F., 2012a. Medieval Dark Earth in an Active Alluvial Setting from the Uffizi Gallery Complex in Florence, Italy. Geoarchaeology 27, 105–122. https://doi.org/10.1002/gea.21403

Nicosia, C., Langohr, R., Mees, F., Arnoldus-Huyzendveld, A., Bruttini, J., Cantini, F., 2012b. Medieval Dark Earth in an Active Alluvial Setting from the Uffizi Gallery ComplNicosia, C., Langohr, R., Mees, F., Arnoldus-Huyzendveld, A., Bruttini, J., Cantini, F., 2012. Medieval Dark Earth in an Active Alluvial Setting from the Uffizi Gallery Complex. Geoarchaeology 27, 105–122. https://doi.org/10.1002/gea.21403

Norman, P., Reader, F.W., 1912. Further discoveries relating to Roman London. Archaeologia 63, 12–344.

Norwood, M.J., Louchouarn, P., Kuo, L.J., Harvey, O.R., 2013. Characterization and biodegradation of water-soluble biomarkers and organic carbon extracted from low temperature chars. Organic Geochemistry 56, 111–119. https://doi.org/10.1016/j.orggeochem.2012.12.008

Oros, D.R., Abas, M.R. bin, Omar, N.Y.M.J., Rahman, N.A., Simoneit, B.R.T., 2006. Identification and emission factors of molecular tracers in organic aerosols from biomass burning: Part 3. Grasses. Applied Geochemistry 21, 919–940. https://doi.org/10.1016/j.apgeochem.2006.01.008

Oros, D.R., Mazurek, M.A., Baham, J.E., Simoneit, B.R.T., 2002. Organic tracers from wild fire residues in soils and rain/river wash-out. Water, Air, and Soil Pollution 137, 203–233. https://doi.org/10.1023/A:1015557301467

Oros, D.R., Simoneit, B.R.T., 2001. Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 2. Deciduous trees. Applied Geochemistry 16, 1545–1565. https://doi.org/10.1016/S0883-2927(01)00022-1

Oros, D.R., Simoneit, B.R.T., 1999. Identification of molecular tracers in organic aerosols from temperate climate vegetation subjected to biomass burning. Aerosol Science and Technology 31, 433–445. https://doi.org/10.1080/027868299303986

Paetsch, L., Mueller, C.W., Rumpel, C., Angst, Š., Wiesheu, A.C., Girardin, C., Ivleva, N.P., Niessner, R., Kögel-Knabner, I., 2017. A multi-technique approach to assess the fate of biochar in soil and to quantify its effect on soil organic matter composition. Organic Geochemistry 112, 177–186. https://doi.org/10.1016/j.orggeochem.2017.06.012

Pandey, S.D., Rocha, L.C., Pereira, G., Deschamps, C., Campos, J.L.E., Falcão, N., Prous, A., Jorio, A., 2020. Properties of carbon particles in archeological and natural Amazon rainforest soils. Catena 194, 104687. https://doi.org/10.1016/j.catena.2020.104687

Partridge, L.G., Djerassi, C., 1977. Mass Spectrometry in Structural and Stereochemical Problems. 250.1 Characteristic Fragmentations of Cholesterol Acetate. Journal of Organic Chemistry 42, 2799–2805.

Partridge, L.G., Midgley, I., Djerassi, C., 1977. Mass Spectrometry in Structural and Stereochemical Problems. 249.1 Elucidation of the Course of the Characteristic Ring D Fragmentation of Unsaturated Steroids2. Journal of the American Chemical Society 99, 7686–7695. https://doi.org/10.1021/ja00465a045

Patten, H. V., Meadows, K.E., Hutton, L.A., Iacobini, J.G., Battistel, D., McKelvey, K., Colburn, A.W., Newton, M.E., MacPherson, J. V., Unwin, P.R., 2012. Electrochemical mapping reveals direct correlation between heterogeneous electron-transfer kinetics and local density of states in diamond electrodes. Angewandte Chemie - International Edition 51, 7002–7006. https://doi.org/10.1002/anie.201203057

Pessenda, L.C.R.., Gouveia, S.E.M.., Aravena, R., 2021. Radiocarbon dating of total soil organic matter and humin fraction and its comparison with 14C ages of fossil charcoal 43, 6.

Peterson, W., 1974. Summary Report of Two Archaeological Sites from North-Eastern Luzon. Archaeology & Physical Anthropology in Oceania 9, 26–35.

Pierce, C., Adams, K.R., Stewart, J.D., 1998. Determining the Fuel Constituents of Ancient Hearth Ash Via ICP-AES Analysis. Journal of Archaeological Science 25, 493–503. https://doi.org/https://doi.org/10.1006/jasc.1997.0252

Piperno, D.R., 2006. Phytoliths: a comprehensive guide for archaeologists and paleoecologists. Rowman Altamira.

Pisani, O., Oros, D.R., Oyo-Ita, O.E., Ekpo, B.O., Jaffé, R., Simoneit, B.R.T., 2013. Biomarkers in surface sediments from the Cross River and estuary system, SE Nigeria: Assessment of organic matter sources of natural and anthropogenic origins. Applied Geochemistry 31, 239–250. https://doi.org/10.1016/j.apgeochem.2013.01.010

Prokof'eva, T. V., Varava, O.A., Sedov, S.N., Kuznetsova, A.M., 2010. Morphological diagnostics of pedogenesis on the anthropogenically transformed floodplains in Moscow. Eurasian Soil Science 43, 368–379. https://doi.org/10.1134/S1064229310040022

Qi, F., Naidu, R., Bolan, N.S., Dong, Z., Yan, Y., Lamb, D., Bucheli, T.D., Choppala, G., Duan, L., Semple, K.T., 2017. Pyrogenic carbon in Australian soils. Science of the Total Environment 586, 849–857. https://doi.org/10.1016/j.scitotenv.2017.02.064

Quéméneur, M., Marty, Y., 1994. Fatty acids and sterols in domestic wastewaters. Water Research 28, 1217–1226. https://doi.org/https://doi.org/10.1016/0043-1354(94)90210-0

Ravindra, K., Sokhi, R., Van Grieken, R., 2008. Atmospheric polycyclic aromatic hydrocarbons: Source attribution, emission factors and regulation. Atmospheric Environment 42, 2895–2921. https://doi.org/10.1016/j.atmosenv.2007.12.010

Reeves, A.D., Patton, D., 2005. Faecal sterols as indicators of sewage contamination in estuarine sediments of the Tay Estuary, Scotland: An extended baseline survey. Hydrology and Earth System Sciences 9, 81–94. https://doi.org/10.5194/hess-9-81-2005

Reimann, C., De Caritat, P., 2000. Intrinsic flaws of element enrichment factors (EFs) in environmental geochemistry. Environmental Science and Technology 34, 5084–5091. https://doi.org/10.1021/es001339o

Ribeiro-Soares, J., Cançado, L.G., Falcão, N.P.S., Martins Ferreira, E.H., Achete, C.A., Jorio, A., 2013. The use of Raman spectroscopy to characterize the carbon materials found in Amazonian anthrosoils. Journal of Raman Spectroscopy 44, 283–289. https://doi.org/10.1002/jrs.4191

Rowley, M.C., Grand, S., Verrecchia, É.P., 2018. Calcium-mediated stabilisation of soil organic carbon. Biogeochemistry 137, 27–49. https://doi.org/10.1007/s10533-017-0410-1

Ruan, Y., Mohtadi, M., Dupont, L.M., Hebbeln, D., van der Kaars, S., Hopmans, E.C., Schouten, S., Hyer, E.J., Schefuß, E., 2020. Interaction of Fire, Vegetation, and Climate in Tropical Ecosystems: A Multiproxy Study Over the Past 22,000 Years. Global Biogeochemical Cycles 34. https://doi.org/10.1029/2020GB006677

Ruddiman, W.F., 2014. Earth's Climate past and future.

Rudnick, R.L., Gao, S., 2013. Composition of the Continental Crust, 2nd ed, Treatise on Geochemistry: Second Edition. Elsevier Ltd. https://doi.org/10.1016/B978-0-08-095975-7.00301-6

Rumpel, C., Alexis, M., Chabbi, A., Chaplot, V., Rasse, D.P., Valentin, C., Mariotti, A., 2006. Black carbon contribution to soil organic matter composition in tropical sloping land under slash and burn agriculture. Geoderma 130, 35–46. https://doi.org/10.1016/j.geoderma.2005.01.007

Rumpel, C., Leifeld, J., Santin, C., Doerr, S., 2015. Movement of biochar in the environment, in: Biochar for Environmental Management. Routledge, p. 18.

Saim, N., Osman, R., Sari Abg Spian, D.R., Jaafar, M.Z., Juahir, H., Abdullah, M.P., Ghani, F.A., 2009. Chemometric approach to validating faecal sterols as source tracer for faecal contamination in water. Water Research 43, 5023–5030. https://doi.org/10.1016/j.watres.2009.08.052

Saito, R., Jorio, A., Souza Filho, A.G., Dresselhaus, M.S., Dresselhaus, G., Pimenta, M.A., 2002. Probing Phonon Dispersion Relations of Graphite by Double Resonance Raman Scattering. Physical Review Letters 88, 4. https://doi.org/10.1103/PhysRevLett.88.027401

Sartori, F., Lal, R., Ebinger, M.H., Eaton, J.A., 2007. Changes in soil carbon and nutrient pools along a chronosequence of poplar plantations in the Columbia Plateau, Oregon, USA. Agriculture, Ecosystems and Environment 122, 325–339. https://doi.org/10.1016/j.agee.2007.01.026

Schmidl, C., Marr, I.L., Caseiro, A., Kotianová, P., Berner, A., Bauer, H., Kasper-Giebl, A., Puxbaum, H., 2008. Chemical characterisation of fine particle emissions from wood stove combustion of common woods growing in mid-European Alpine regions. Atmospheric Environment 42, 126–141. https://doi.org/10.1016/j.atmosenv.2007.09.028

Schmidt, M.W.I., Noack, A.G., 2000. Black carbon in soils and sediments: Analysis, distribution, implications, and current

challenges 14, 777–793.

Schreuder, L.T., Hopmans, E.C., Castaneda, I.S., Schefuss, E., Sinninghe Damste, J.S., Schouten, S., 2018. Biomass burning in the Sahara/Sahel region in Northwest Africa and interaction with climate, vegetation and humans, in: EGU General Assembly Conference Abstracts. p. 2955.

Schüpbach, S., Kirchgeorg, T., Colombaroli, D., Beffa, G., Radaelli, M., Kehrwald, N.M., Barbante, C., 2015. Combining charcoal sediment and molecular markers to infer a Holocene fire history in the Maya Lowlands of Petén, Guatemala. Quaternary Science Reviews 115, 123–131. https://doi.org/10.1016/j.quascirev.2015.03.004

Schwan, J., Ulrich, S., Batori, V., Ehrhardt, H., Silva, S.R.P., 1996. Raman spectroscopy on amorphous carbon films. Journal of Applied Physics 80, 440–447. https://doi.org/10.1063/1.362745

Shafizadeh, F., 1984. The chemistry of pyrolysis and combustion, in: Rowell, R. (Ed.), Chemistry of Solid Wood. Adv. Chem. Series 207, pp. 489–529.

Shah, V.G., Hugh Dunstan, R., Geary, P.M., Coombes, P., Roberts, T.K., Von Nagy-Felsobuki, E., 2007. Evaluating potential applications of faecal sterols in distinguishing sources of faecal contamination from mixed faecal samples. Water Research 41, 3691–3700. https://doi.org/10.1016/j.watres.2007.04.006

Sheka, E.F., Golubev, Y.A., Popova, N.A., 2020. Graphene domain signature of raman spectra of sp2 amorphous carbons. Nanomaterials 10, 1–22. https://doi.org/10.3390/nano10102021

Shi, H.J., Wang, X.J., Zhao, Y.J., Xu, M.G., Li, D.W., Guo, Y., 2017. Relationship between soil inorganic carbon and organic carbon in the wheat-maize cropland of the North China Plain. Plant and Soil 418, 423–436. https://doi.org/10.1007/s11104-017-3310-1

Silliman, J.E., Meyers, P.A., Bourbonniere, R.A., 1996. Record of postglacial organic matter delivery and burial in sediments of Lake Ontario. Organic Geochemistry 24, 463–472. https://doi.org/10.1016/0146-6380(96)00041-1

Silva, L.C.R., Corrêa, R.S., Wright, J.L., Bomfim, B., Hendricks, L., Gavin, D.G., Muniz, A.W., Martins, G.C., Motta, A.C.V., Barbosa, J.Z., Melo, V. de F., Young, S.D., Broadley, M.R., Santos, R.V., 2021. A new hypothesis for the origin of Amazonian Dark Earths. Nature Communications 12, 1–11. https://doi.org/10.1038/s41467-020-20184-2

Silva, T.R., Lopes, S.R.P., Spörl, G., Knoppers, B.A., Azevedo, D.A., 2012. Source characterization using molecular distribution and stable carbon isotopic composition of n-alkanes in sediment cores from the tropical Mundaú-Manguaba estuarine-lagoon system, Brazil. Organic Geochemistry 53, 25–33. https://doi.org/10.1016/j.orggeochem.2012.05.009

Simoneit, B.R.T., 2002. Biomass burning - A review of organic tracers for smoke from incomplete combustion, Applied Geochemistry. https://doi.org/10.1016/S0883-2927(01)00061-0

Simoneit, B.R.T., Elias, V.O., 2000. Organic tracers from biomass burning in atmospheric particulate matter over the ocean. Marine Chemistry 69, 301–312. https://doi.org/10.1016/S0304-4203(00)00008-6

Simoneit, B.R.T., Schauer, J.J., Nolte, C.G., Oros, D.R., Elias, V.O., Fraser, M.P., Rogge, W.F., Cass, G.R., 1999. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. Atmospheric Environment 33, 1–10.

Smidt, E., Tintner, J., Klemm, S., Scholz, U., 2017. FT-IR spectral and thermal characterization of ancient charcoals - A tool to support archeological and historical data interpretation. Quaternary International 457, 43–49. https://doi.org/10.1016/j.quaint.2016.11.031

Smidt, E., Tintner, J., Nelle, O., Oliveira, R.R., Patzlaff, R., Novotny, E.H., Klemm, S., 2020. Infrared spectroscopy refines chronological assessment, depositional environment and pyrolysis conditions of archeological charcoals. Scientific Reports 10, 1–11. https://doi.org/10.1038/s41598-020-69445-6

Smith, M.W., Dallmeyer, I., Johnson, T.J., Brauer, C.S., McEwen, J.S., Espinal, J.F., Garcia-Perez, M., 2016. Structural analysis of char by Raman spectroscopy: Improving band assignments through computational calculations from first principles. Carbon 100, 678–692. https://doi.org/10.1016/j.carbon.2016.01.031

Song, Y.Y., Lu, Y., 2015. Decision tree methods: applications for classification and prediction. Shanghai Archives of Psychiatry 27, 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

Steiner, C., Teixeira, W.G., Lehmann, J., Nehls, T., De MacÊdo, J.L.V., Blum, W.E.H., Zech, W., 2007. Long term effects of manure, charcoal and mineral fertilization on crop production and fertility on a highly weathered Central Amazonian upland soil. Plant and Soil 291, 275–290. https://doi.org/10.1007/s11104-007-9193-9

Stoops, G., 2003. Guidelines for analysis and description of soil and regolith thin sections. Soil Science Society of America, Madison, WI.

Sukopp, H., 1979. The soil, flora, and vegetation of Berlin's waste lands.

Sutherland, R.A., 2000. Bed sediment-associated trace metals in an urban stream, Oahu, Hawaii. Environmental Geology 39, 611–627. https://doi.org/10.1007/s002540050473

Tan, Z., Wu, C., Han, Y., Zhang, Y., Mao, L., Li, D., Liu, L., Su, G., Yan, T., 2020. Fire history and human activity revealed through poly cyclic aromatic hydrocarbon (PAH) records at archaeological sites in the middle reaches of the Yellow River drainage basin, China. Palaeogeography, Palaeoclimatology, Palaeoecology 560, 110015. https://doi.org/10.1016/j.palaeo.2020.110015

Tanaka, C., Watanuki, M., Nazaki, Y., 1983. Effect of soybean protein on coprostanol production and cholesterol metabolism in cholesterol-fed rats. Journal of nutritional science and vitaminology 29, 447–454.

Tharwat, A., Gaber, T., Ibrahim, A., Hassanien, A.E., 2017. Linear discriminant analysis: A detailed tutorial. AI Communications 30, 169–190. https://doi.org/10.3233/AIC-170729

Theurer, T., Muirhead, D.K., Jolley, D., Mauquoy, D., 2021. The applicability of Raman spectroscopy in the assessment of palaeowildfire intensity. Palaeogeography, Palaeoclimatology, Palaeoecology 570, 110363. https://doi.org/10.1016/j.palaeo.2021.110363

Thomsen, C., Reich, S., 2000. Double resonant raman scattering in graphite. Physical Review Letters 85, 5214–5217. https://doi.org/10.1103/PhysRevLett.85.5214

Trusovas, R., Račiukaitis, G., Niaura, G., Barkauskas, J., Valušis, G., Pauliukaite, R., 2016. Recent Advances in Laser Utilization in the Chemical Modification of Graphene Oxide and Its Applications. Advanced Optical Materials 4, 37–65. https://doi.org/10.1002/adom.201500469

Tuinstra, F., Koenig, J., 1970. Raman Spectrum of Graphite. Journal of Chemical Physics 53, 1126–1130. https://doi.org/10.1063/1.1674108

Tyagi, P., Edwards, D.R., Coyne, M.S., 2008. Use of sterol and bile acid biomarkers to identify domesticated animal sources of fecal pollution. Water, Air, and Soil Pollution 187, 263–274. https://doi.org/10.1007/s11270-007-9514-x

Tyagi, P., Edwards, D.R., Coyne, M.S., 2007. Use of selected chemical markers in combination with a multiple regression model to assess the contribution of domesticated animal sources of fecal pollution in the environment. Chemosphere 69, 1617–1624. https://doi.org/10.1016/j.chemosphere.2007.05.039

van der Plicht, J., Streurman, H.J., van Mourik, J.M., 2019. Radiocarbon dating of soil archives, Developments in Quaternary Science. https://doi.org/10.1016/B978-0-444-64108-3.00003-3

van Mourik, J.M., Seijmonsbergen, A.C., Slotboom, R.T., Wallinga, J., 2012. Impact of human land use on soils and landforms in cultural landscapes on aeolian sandy substrates (Maashorst, SE-Netherlands). Quaternary International 265, 74–89. https://doi.org/10.1016/j.quaint.2011.06.053

Venkatesan, M.I., Kaplan, I.R., 1990. Sedimentary Coprostanol as an Index of Sewage Addition in Santa Monica Basin, Southern California. Environmental Science and Technology 24, 208–214. https://doi.org/10.1021/es00072a009

Venkatesan, M.I., Santiago, C.A., 1989. Sterols in ocean sediments: novel tracers to examine habitats of cetaceans, pinnipeds, penguins and humans. Marine Biology 102, 431–437. https://doi.org/10.1007/BF00438343

Vittori Antisari, L., Cremonini, S., Desantis, P., Calastri, C., Vianello, G., 2013. Chemical characterisation of anthro-technosols from Bronze to Middle Age in Bologna (Italy). Journal of Archaeological Science 40, 3660–3671. https://doi.org/10.1016/j.jas.2013.04.023

Vrydaghs, L., Ball, T.B., Devos, Y., 2016. Beyond redundancy and multiplicity. Integrating phytolith analysis and micromorphology to the study of Brussels Dark Earth. Journal of Archaeological Science 68, 79–88. https://doi.org/10.1016/j.jas.2015.09.004

Vrydaghs, L., Devos, Y., Charruadas, P., Cummings, L.S., Degraeve, A., 2019. Agricultural Activities in the 10th–13th Century CE in Brussels (Belgium): An Interdisciplinary Approach. Agricultural and Pastoral Landscapes in Pre-Industrial Society 221–234. https://doi.org/10.2307/j.ctvh1dswm.22

Wang, Y., Alsmeyer, D.C., McCreery, R.L., 1990. Raman Spectroscopy of Carbon Materials: Structural Basis of Observed Spectra.

Chemistry of Materials 2, 557–563. https://doi.org/10.1021/cm00011a018

Wang, Z., Fingas, M., Sigouin, L., Landriault, M., Li, K., Lambert, P., Turpin, R., Campagna, P., Mullin, J., 1998. Quantitative characterization of PAHs in burn residue and soot samples and differentiation of pyrogenic PAHs from petrogenic PAHs - The 1994 mobile burn study. Proceedings - 21st Arctic and Marine Oil Spill Program Technical Seminar 1998 33, 673–703.

Ward, T.J., Hamilton, R.F., Dixon, R.W., Paulsen, M., Simpson, C.D., 2006. Characterization and evaluation of smoke tracers in PM: Results from the 2003 Montana wildfire season. Atmospheric Environment 40, 7005–7017. https://doi.org/10.1016/j.atmosenv.2006.06.034

Wells, E.C., Terry, R.E., Parnell, J.J., Hardin, P.J., Jackson, M.W., Houston, S.D., 2000. Chemical Analyses of Ancient Anthrosols in Residential Areas at Piedras Negras, Guatemala. Journal of Archaeological Science 27, 449–462. https://doi.org/https://doi.org/10.1006/jasc.1999.0490

Whelton, H.L., Hammann, S., Cramp, L.J.E., Dunne, J., Roffet-Salque, M., Evershed, R.P., 2021. A call for caution in the analysis of lipids and other small biomolecules from archaeological contexts. Journal of Archaeological Science 132. https://doi.org/10.1016/j.jas.2021.105397

Whitlock, C., Larsen, C., 2002. Charcoal as a Fire Proxy 75–97. https://doi.org/10.1007/0-306-47668-1_5

Wiedner, K., Schneeweiß, J., Dippold, M.A., Glaser, B., 2015. Anthropogenic Dark Earth in Northern Germany - The Nordic Analogue to terra preta de Índio in Amazonia. Catena 132, 114–125. https://doi.org/10.1016/j.catena.2014.10.024

Wilcke, W., 2007. Global patterns of polycyclic aromatic hydrocarbons (PAHs) in soil. Geoderma 141, 157–166. https://doi.org/10.1016/j.geoderma.2007.07.007

Wouters, B., 2020. A Biographical Approach to Urban Communities from a - Geoarchaeo - logical Perspective. High-definition applications and Case Studies. Journal of Urban Archaeology 2, 85–101.

Wouters, B., 2016. Geoarchaeological and micromorphological approaches to the formation and biographies of Early Medieval towns in Northwest Europe.

Wouters, B., Devos, Y., Milek, K., Vrydaghs, L., Bartholomieux, B., Tys, D., Moolhuizen, C., van Asch, N., 2017. Medieval markets: A soil micromorphological and archaeobotanical study of the urban stratigraphy of Lier (Belgium). Quaternary International 460, 48–64. https://doi.org/10.1016/j.quaint.2017.03.002

Wouters, B., Devos, Y., Vrydaghs, L., Ball, T., De Winter, N., Reygel, P., 2019. An integrated micromorphological and phytolith study of urban soils and sediments from the Gallo-Roman town Atuatuca Tungrorum, Belgium. Geoarchaeology 34, 448–466. https://doi.org/10.1002/gea.21722

Wu, J., Rees, P., Storrer, S., Alderisio, K., Dorner, S., 2009. Fate and transport modeling of potential pathogens: The contribution from sediments. Journal of the American Water Resources Association 45, 35–44. https://doi.org/10.1111/j.1752-1688.2008.00287.x

Wyllie, S.G., Amos, B.A., Tökés, L., 1977. Electron Impact Induced Fragmentation of Cholesterol and Related C-5 Unsaturated Steroids. Journal of Organic Chemistry 42, 725–732. https://doi.org/10.1021/jo00424a033

Xanthopoulos, P., Pardalos, P., Trafalis, T., 2012. Robust Data Mining. Briefs in Optimization 1–52.

Xie, H., Zhuang, X., Bai, Z., Qi, H., Zhang, H., 2006. Isolation of levoglucosan-assimilating microorganisms from soil and an investigation of their levoglucosan kinases. World Journal of Microbiology and Biotechnology 22, 887–892.

Xie, S., Chen, F., Wang, Z., Wang, H., Gu, Y., Huang, Y., 2003. Lipid distributions in loess-paleosol sequences from northwest China. Organic Geochemistry 34, 1071–1079. https://doi.org/10.1016/S0146-6380(03)00083-4

Yu, X.Y., Ying, G.G., Kookana, R.S., 2006. Sorption and desorption behaviors of diuron in soils amended with charcoal. Journal of Agricultural and Food Chemistry 54, 8545–8550. https://doi.org/10.1021/jf061354y

Yule, B., 1990. The “dark earth” and late Roman London. Antiquity 64, 620–628.

Yunker, M.B., Macdonald, R.W., Vingarzan, R., Mitchell, H., Goyette, D., Sylvestre, S., 2002. PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition. Organic Geochemistry 33, 489–515.

Zech, M., Buggle, B., Leiber, K., Marković, S., Glaser, B., Hambach, U., Huwe, B., Stevens, T., Sümegi, P., Wiesenberg, G., Zöller, L., 2010. Reconstructing Quaternary vegetation history in the Carpathian Basin, SE-Europe, using n-alkane biomarkers as molecular fossils: Problems and possible solutions, potential and limitations. E&G Quaternary Science Journal 58, 148–155. https://doi.org/10.3285/eg.58.2.03

Zech, M., Krause, T., Meszner, S., Faust, D., 2013. Incorrect when uncorrected: Reconstructing vegetation history using n-alkane biomarkers in loess-paleosol sequences - A case study from the Saxonian loess region, Germany. Quaternary International 296, 108–116. https://doi.org/10.1016/j.quaint.2012.01.023

Zethof, J.H.T., Leue, M., Vogel, C., Stoner, S.W., Kalbitz, K., 2019. Identifying and quantifying geogenic organic carbon in soils - The case of graphite. Soil 5, 383–398. https://doi.org/10.5194/soil-5-383-2019

Zhang, G.L., Burghardt, W., Lu, Y., Gong, Z.T., 2001. Phosphorus-enriched soils of urban and suburban Nanjing and their effect on groundwater phosphorus. Journal of Plant Nutrition and Soil Science 164, 295–301. https://doi.org/10.1002/1522-2624(200106)164:3<295::AID-JPLN295>3.0.CO;2-T

Zhang, J., Cornelia, M.N., Wang, M., Cao, Z., Luo, X., Wong, M., Chen, W., 2013. Change of PAHs with evolution of paddy soils from prehistoric to present over the last six millennia in the Yangtze River Delta region, China. Science of the Total Environment 449, 328–335. https://doi.org/10.1016/j.scitotenv.2013.01.084

Zhang, X., Qiao, X.F., Shi, W., Wu, J. Bin, Jiang, D.S., Tan, P.H., 2015. Phonon and Raman scattering of two-dimensional transition metal dichalcogenides from monolayer, multilayer to bulk material. Chemical Society Reviews 44, 2757–2785. https://doi.org/10.1039/c4cs00282b

Zhao, Z., 2011. New archaeobotanic data for the study of the origins of agriculture in China. Current Anthropology 52, 295–306. https://doi.org/10.1086/659308

Zickler, G.A., Smarsly, B., Gierlinger, N., Peterlik, H., Paris, O., 2006. A reconsideration of the relationship between the crystallite size La of carbons determined by X-ray diffraction and Raman spectroscopy. Carbon 44, 3239–3246. https://doi.org/10.1016/j.carbon.2006.06.029

Zorzin, R., 2021. Geologia e Paleontologia dell ' area collinare di Verona.

Zou, S., Li, R., Xie, S., Zhu, J., Wang, X., Huang, J., 2010. Paleofire indicated by polycyclic aromatic hydrocarbons in soil of Jinluojia archaeological site, Hubei, China. Journal of Earth Science 21, 247–256. https://doi.org/10.1007/s12583-010-0089-x

Acknowledgements

This thesis is the result of work and study during the last three and a half years. This period was exciting and challenging, and it would not be the same without the contribution of many people, which inevitably I will not be able to thank; it is necessary, however, for me to try.

First and foremost, I want to express my deepest gratitude to my supervisor, Professor Dario Battistel for his prompt presence. His way of approaching any scientific question, changing perspective and point of view, represents for me a model and the significance of making research, and made this journey inspiring and stimulating.

I would also like to thank my co-supervisor, professor Lavinia de Ferri for her support during my abroad period in Norway and her carefulness in revising this thesis. I would also like to express my gratitude to Professor Cristiano Nicosia, for having made available his expertise and the samples of this research.

I also want to thank Professor Giulio Pojana and Federica Agnoletto for their work on the Raman analysis.

I would also to give thanks to professor Andrea R Brunelle and Doctor Yannick Devos for having carefully read this thesis and for their precious suggestions and revisions.

This period would not be the same without the people that work, study, and live on the scientific campus of Via Torino, especially the Delta and Epsilon buildings, I hope you would not get upset if I don't mention you all. Just Elena: I hope this time the commas are not so randomly placed.

I want to thank my friends Marietta and Edo, for being always here, and Fernando, for our endless complaints and his help with the title.

Lastly, I want to thank Fabrizio, the love of my life, and my family: Anna, Mamma e Papà, for their love and their strength.