Distribution of *n*-alkyl lipids in particle size fractions of soil with different vegetation cover and under similar environmental conditions: a case study from Lower Gangetic Plain

MS project report submitted in fulfillment of the requirements

for the degree of

Master of Science

By

Twismary Kharphuli

16MS066

Under the supervision of

Professor Prasanta Sanyal

to the

Department of Earth Sciences

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH KOLKATA
MOHANPUR, NADIA 741246

July 23, 2021

Declaration by the student

Date: 23rd July 2021

I, **Miss Twismary Kharphuli**, Registration Number **16MS066** dated **29th July 2016**, a student of **Department of Earth Sciences** of the BS-MS Program of **IISER Kolkata**, hereby declare that this MS project report is composed entirely by myself, solely the result of my own work and not submitted for any other degree or professional qualification. To the best of my knowledge, it does not contain materials previously published or written by any other person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at IISER Kolkata or any other educational institution except where due acknowledgment is made in the report.

I hereby grant permission to IISER Kolkata to store the project report in a database which can be accessed by others.

Miss Twismary Kharphuli

Department of Earth Sciences

Indian Institute of Science Education and Research Kolkata

Mohanpur 741246, West Bengal, India

Certificate from the supervisor

This is to certify that the research work encapsulated in the MS project report entitled "**Distribution of *n*-alkyl lipids in particle size fractions of soil with different vegetation cover and under similar environmental conditions: a case study from Lower Gangetic Plain"** submitted by **Miss Twismary Kharphuli 16MS066,** a student of Department of Earth Sciences of the BS-MS Program of IISER Kolkata under my supervision and guidance is based upon her own work to be submitted for the requirements for awarding the **BS-MS Dual Degree** by Indian Institute of Science Education and Research Kolkata. The content of this project has not been submitted elsewhere for the award of any academics and a professional degree.

Date: 21st July 2021

Signature

Dr Prasanta Sanyal

Professor

Department of Earth Sciences

Indian Institute of Science Education and Research Kolkata

Mohanpur 741246, West Bengal, India

Acknowledgement

Over the last year that has taken me to complete my dissertation, I am eternally grateful to the following people who have guided me and provided valuable inputs to my Master's project.

First and foremost, I would like to acknowledge and render my warmest thanks to my supervisor, **Prof. Prasanta Sanyal,** for providing me with the opportunity to work in his laboratory. His scholarly inputs, scientific discussion, useful comments, remarks, and engagement throughout the learning process has helped me to steer the project in the right direction.

I am indebted and wish to express my sincere thanks to **Dr. Biswajit Roy** for the valuable input, who contributed in many discussions that help to shape this project. I would like to thank you very much for your assistance and dedicated involvement in every step throughout the process.

My gratitude extends to **Mr. Mahesh Ghosh** for his mentorship and technical support in the laboratory.

These acknowledgements would not be complete without mentioning my friends and all the members of Stable Isotopes Laboratory IISER Kolkata. Their unbiased opinion and constructive comments helped me during my time at the laboratory. The experience I gained from my time in the lab will surely assist me in my future endeavors.

Most importantly, none of this could have happened without my parents. Without their tremendous understanding, strong and consistent support it would be impossible for me to complete my study. This dissertation stands as a testament to their unconditional love, patience, everlasting encouragement and prayers.

Abstract

Chemical stabilization of organic matter (OM) with clay minerals and protection by occlusion within microaggregates are the important mechanisms responsible for the storage of different OM pools in the soils. Soil lipids constitute an important fraction of OM and are found to be resistant against degradation from post-depositional processes. The hydrophobic properties and high reactivity with polyvalent cations makes the lipids more stable in comparison to other OM pools. Such properties of lipid biomarkers are widely explored to extract the information about different sources and the extent of degradation and/or preservation mechanisms of OM in sediments. Although the lipids are stable across different ecosystems, it has been observed that the difference in vegetation composition, grain size distribution, climatic conditions as well as the mineralogical association can significantly influence the distribution and stabilization of lipid pools in the soil/sediments. However, previous studies emphasizing on the distribution of *n*-alkyl lipids in soil across different vegetation growth did not include the mineralogical and textural influences on these compounds in the soil. Additionally, relative abundance and composition among different particle size fractions were mostly focused on a particular vegetation composition without any possibility of comparison of lipids produced from different vegetation types and their behavior under similar climatic conditions. The present study emphasizes the importance of vegetation cover and also the micro-environmental variations in controlling the distribution of *n*-alkyl lipids in the soil from lower Gangetic Plain. This region is one of the largest fluvial settings and biodiversity rich ecosystems dominated by grasses and trees and it provides an opportunity to study the respective contribution of the various growth forms to the soil lipids where the soil textural composition and climatic condition are similar. *n*-Alkyl lipids were analyzed from the whole soils and also across five different particle size fractions (250-125μm, 125-63μm, 63-32μm, 32-20μm and <20μm) separated out from forest, grassland and mixed system that comprise of shrubs and grasses. The concentration and distribution pattern of *n*-alkanes and *n*-alkanoic acids in whole soils under different vegetation cover was different. Highest total *n*-alkanes abundance was found in forest (5.09μg/g), lowest in mixed system (1.06μg/g) and intermediate in grassland (3.23μg/g). Deviation in *n*-alkanes concentration in these soils depends on the chain length distribution, growth habits and physiology of different vegetation types, all of which influence *n*-alkanes signals in soils. Total *n*-alkanoic acid abundance was found to be higher in grassland (6.90 μg/g), comparable in forest (5.09μg/g) and lowest in mixed (0.82μg/g) systems. A large contribution from the even carbon-number low molecular weight (LMW) compounds (C_{12}-C_{20}) is the reason for the observed higher total *n*-alkanoic acid concentration in grassland soil. High molecular weight compounds (HMW) abundance of both *n*-alkyl compounds, are highest in forest soil (4.56μg/g and 29.93μg/g) compared to grassland

(2.55μg/g and 23.90μg/g) and mixed system (0.7μg/g and 2.40μg/g) systems respectively. Better preservation of long-chain *n*-alkanes and *n*-alkanoic acids in soils under forests could result from a thick layer of litter assemblages accumulated in the topsoil surfaces that prevent the infiltration of air and water leading to less microbial communities and possibly result in the preservation of long-chain compounds in the soil. LMW abundance of *n*-alkanes and *n*-alkanoic acids are relatively higher in grassland soil (0.68μg/g and 45.16μg/g) than forest (0.52μg/g and 21.03μg/g) and mixed system soil (0.35μg/g and 5.86μg/g). This is attributed to the specificity of grassland where the roots play an important role in proliferation of microbial growth that influences the degradation of HMW to LMW compounds. To evaluate the *n*-alkyl lipid distributions in these soils that can derive both from plants and microorganisms, we separated five different particle size fractions of soils collected from these three different ecosystems. Different size fractions across different soil systems showed wide variation in their total *n*-alkyl concentrations, odd/even predominance and in LMW/HMW ratio compared to the values observed in the whole soils. Despite differences in vegetation cover, the highest total *n*-alkyl lipid was observed in the finer fractions, which denotes the adsorption of lipids to clay mineral surfaces and/or stabilized by occluding into the microaggregates. Long-chain odd *n*-alkanes and even *n*-alkanoic acid compounds are abundant in sand fractions. This indicates the presence of plant material in sand fractions that is attributed to the dominant minerals quartz present in sand-size fractions. Quartz minerals have less capability to hold the OM because of thin clay coating compared to the phyllosilicates minerals which predominate in the finer fractions. Silt fractions have the lowest HMW *n*-alkyl compounds. Substantial degradation of long chain compounds could have been attributed to the higher microbial biomass found in silt size fractions. High abundance of both HMW and LMW *n*-alkyl compounds with lower odd/even predominance in finer fractions is thought to be attributed to degradation of long chain to short chain compounds, along with a significant contribution from the microbial communities in the lower chain compounds. Molecular indices deriving from these two compounds yield a large potential to elucidate the sources of *n*-alkanes and *n*-alkanoic acids entering the whole soil, and in combination with particle size fractions they enable the identification of incorporation and preservation pathways of lipids in soil. Long-chain CPI of *n*-alkanes ($CPI_{alk(23-33)}$) is 7.0, 5.2 and 2.1; *n*-alkanoic acids ($CPI_{acid(22-32)}$) is 2.4, 2.6 and 3.0 in forest, grassland and mixed vegetation whole soils respectively reflect that the long-chain fractions were largely derived from terrestrial higher plants and not significantly altered by diagenesis. CPI of *n*-alkanes ($CPI_{alk(23-33)}$) indicates a more significant variation in HMW carbon number predominance between different systems compared to the *n*-alkanoic acids ($CPI_{acid(22-32)}$). The $CPI_{alk(23-33)}$ values in particle size fractions of forest, grassland and mixed vegetation soil shows a systematic decreasing trend with decreasing size

fractions of 30-60% from coarser to finer size fractions. This confirmed the plant derived *n*-alkanes dominated in sand and microbial modified fractions in the finer fractions. $CPI_{acid(22-32)}$ values of *n*-alkanoic acids show minimal variation among particle size fractions with a decreasing pattern of 10-20% from sand size to finer fractions, an indication of long-term turn over *n*-alkanoic acids accumulated in the finer fractions. A possible protection mechanism points to the distinctive structure of *n*-alkanoic acids with functional groups that interact with the clay minerals. The ACL of long chain *n*-alkanes ($ACL_{alk(25-33)}$) in forest is 29.8, grassland is 30.6 and mixed system is 29.4. Lipids in leaves derived from plants in forests on average have shorter chain lengths than lipids in leaves from grasslands. The ACL of long chain *n*-alkanoic acids ($ACL_{acid(22-34)}$) in whole soils (27.2-27.7) are roughly constant from all the systems. The similarities in $ACL_{acid(22-34)}$ of long chain *n*-alkanoic acids across the systems rule out its use as a vegetation types indicator, while the difference in *n*-alkane ACL in these systems possibly make it a potentially good indicator for paleoecological application. However, across the size fractions, $ACL_{alk(25-33)}$ and $ACL_{acid(22-34)}$ values calculated from the *n*-alkanes and *n*-alkanoic acids did show a slight but insignificant variation with decreasing grain size (decrease of ~1-2% in $ACL_{alk(25-33)}$ and ~3% in $ACL_{acid(22-34)}$), suggesting a role of post-depositional processes in controlling the dominant chain length distribution of *n*-alkyl compounds. The variability of ACL values of *n*-alkyl lipids across different size fractions thus raises possible concerns in the use of these proxies for paleoecological studies. Low odd over predominance ($OEP_{(27-33)}$) in grassland (6.0) and mixed systems (2.5) whole soil compared to forest (7.3) indicating that *n*-alkanes degradation is higher in grassland and mixed systems. The presence of higher fine root density allows more infiltration of water and air, promoting higher proliferation of microbial communities and possibly allowing increased re-working of plant material, contributed by the grassland and mixed systems. Preservation in forest soils is probably biased by the thick litter assemblages in the forest soil leading to anoxic conditions for higher preservation of plant derived OM materials. Microbial activity is favorable under oxic conditions and abruptly decreased under anoxic conditions. In size fractions, a significant decrease (~40%) in $OEP_{(27-33)}$ from sand to finer fraction is observed in all the systems. $OEP_{(27-33)}$ resembles the decrease in the $CPI_{alk(23-33)}$ for the *n*-alkanes. High $OEP_{(27-33)}$ are characteristic of fresh plant material in sand size, whereas low $OEP_{(27-33)}$ in finer fractions indicate modification of *n*-alkanes during early diagenesis. Even over odd predominance ($EOP_{(20-30)}$) in forest (3.5) grassland (3.6) and mixed vegetation soil (3.9) imply that despite considerable difference in *n*-alkanoic acid abundance in these systems, Whereas, in size fractions, all the three systems show a slight decreasing trend of (~5-8%) in the $EOP_{(20-30)}$ from sand size to finer fractions resembles the minimal decrease in the $CPI_{acid(22-32)}$ for the *n*-alkanoic acids, which ascribe to limited modification of long chain *n*-alkanoic acids in coarser fractions and higher degradation in finer fractions. Thus, we demonstrated at the molecular level that

the incorporation of individual C sources varies with the size fractions. Based on molecular abundances and distribution in *n*-alkyl lipids across five different size fractions we evaluated several molecular proxies, which help to elucidate the residence time of these compounds and their importance in global climate studies. Our data from this study suggest that, for these soils, not only the inherent recalcitrance of these compounds, but the vegetation source and micro-environmental variations is also responsible for variation in preservation, transformation and turnover rates of lipid pools.

Table of contents

Contents

List of Figures

List of Tables

List of Equations

Plan of the project report

This report has six chapters. In chapter 1 a brief introduction on how different grain size OM distribution study is better than whole soil study, the process/factors that can influence the distribution of *n*-alkyl lipids and what was not taken in consideration in previous studies. In Chapter 2, we describe the vegetation, climatic and geomorphologic characteristics of the study area from where the sampling of the soil was carried out. In Chapter 3, we briefly discuss the sampling of soil, sieving analysis, analysis of soil textural and mineralogical characteristics and quantification of *n*-alkyl compounds in GC-MS. The results are provided as tabulated data and bar diagram in Chapter 4. Chapter 5 deals with the observations, interpretations and discussion of the result. Finally, we conclude in Chapter 6.

CHAPTER 1

Introduction

Soil organic matter (OM) is known to be the largest pool of carbon on the Earth surface. The biogeochemical cycling of carbon between the soil and the atmosphere mostly depends on the duration of OM storage in the soil. Differences in storage of carbon can vary between a few years to several decades and up to millennia, depending on the pools of OM in the soil. Understanding the behavior of stable fraction of OM in the soil helps in estimating the residence time of oldest OM pools (Eswaran et al., 1993; Alperin et al., 1993; Amundson 2001; Queanea et al., 2004). However, a significant amount of carbon presently stored in the soil as the stable OM fractions can go back into the atmosphere, as a result of intrinsic and extrinsic factors such as climate, weathering, soil substrates, including biological interactions and vegetation compositions. Thus, identification of possible factors impacting the turnover rate of stable OM fractions would help in understanding the pathways and storage of OM pools.

One important class of stable organic pools in soil OM is the lipids (Jansen and Wiesenberg, 2017). Lipids are a group of *n*-alkyl compounds including *n*-alkanes, *n*-alkanoic acids and *n*-alcohols (Eglinton et al., 1996). Soil lipids constitute an important fraction of soil OM and its abundance can reach up to 42 wt% of the total soil organic carbon (Preston et al., 1987; Queneau et al., 2004). *n*-Alkyl compounds such as *n*-alkanes and *n*-alkanoic acid are characterized by their alkyl chains that promote formation of hydrophobic zones, while the carboxyl groups of *n*-alkanoic acids are capable of interacting with the charged minerals surfaces via ligand exchange and polyvalent cation bridges (Lorenz et al., 2007; Yang et al 2020). Such intrinsic properties of *n*-alkyl lipids are of major importance to soil aggregate stability and water retention in the soil (Jambu et al., 1978). Their water insolubility and resistance to biodegradation allows relatively higher preservation to degradation from post-depositional modifications. Such a robust nature of these *n*-alkyl compounds is useful in providing extensive information about the different OM sources and their degradation and preservation mechanisms in the soil.

Chemical adsorption to clay mineral surfaces through the formation of organo-mineral associations by the functional groups of *n*-alkyl molecules and their occlusion within soil microaggregates allows persistency of *n*-alkanes and *n*-alkanoic acids in the soils. Although lipid compounds are bound to minerals, the availability of such compounds varies with grain size. The analysis of *n*-alkyl lipids in size fractions is more useful than the analysis of the entire soil. Lipid concentrations in whole soil can

be largely occluded within soil aggregates, making it difficult to quantify the absolute amount of specific *n*-alkyl compounds present in the soil (Figure a). Separation of soil according to their particle size (sand, silt and clay-size) yields organo-mineral fractions that are different in terms of the amount of the organic compound associated with the mineral grains and the turnover-rates (Christensen, 1996; Queneau et al., 2004). Therefore, lipid analysis across different size fractions may provide more precise information about the long term preservation of lipids through physical and chemical stabilization across the grain size (Thompson and Eglinton, 1978; Six et al., 2002; Lutzow et al., 2006; Lorenz et al., 2007).

Figure a: *Schematic diagram showing the organic matter occluded within the soil aggregates associated with minerals. Separation of soil according to their size yields organo-mineral fractions with distinctly different properties*

Despite the fact that studies on the specific interaction of OM with minerals associated with specific grain size provide detailed information on the stabilization and preservation of lipid compounds with long and short turnover rates, the incorporation of OM from different OM sources and their interaction and preservation pathways is not yet known across the grain size fractions. *n*-Alkyl compounds can be originated from various sources, such as terrestrial higher plants, bacteria and algae (Eglinton and Hamilton, 1967). Long-chain *n*-alkyl compounds are found mainly in leaf waxes of terrestrial higher plants (Eglinton et al., 1962; Rielley et al., 1991) while short chains are produced mostly by microorganisms like fungi, algae and bacteria and also higher plants (Eglinton and Calvin, 1967; Matsuda and Koyama, 1977; Simoneit, 1978). Leaf litter is considered to be a primary source of *n*-alkanes and *n*-alkanoic acids in topsoil (Gamarra and Kahmen, 2015; Angst et al., 2016; Jansen and Wiesenberg 2017). These *n*-alkyl compound ratios in the soils often can be correlated with the plant community of the aboveground vegetation. Leaves constitute an important source of OM in the soil;

however, their abundance in the soil can be influenced by different vegetation types (Quideau et al., 2001).

The type of the overlying vegetation (e.g. trees, shrubs and grasses) strongly determine the content and composition of lipids in the soil (Dinel et al., 1990). Grasses as well as trees are all characterized by different qualitative contribution to soil OM due to their differences in lipid biosynthesis (Rommerskirchen et al., 2006; Wiesenberg and Schwark, 2006; Mueller et al., 2013; Friemuth et al., 2017). Moreover, in grasslands the underground plant biomass (e.g. roots) is greater than aboveground biomass compared to forest systems. As a consequence, it can be assumed that the fraction of OM input to the soil belowground is larger in grassland than in forest where most input occurs in the form of surface leaf litter that accumulated in the topsoil. Consistent with the relative important of roots, microbial activity appears to be greater in grassland than in other ecosystems (Cheng et al., 1990; Mason et al., 2005). Thus, soil under different vegetation cover appears to profoundly affect the contribution of plant and microbial compounds to stabilized OM. Post-depositional modification by microorganisms can also impact molecular abundance and the distribution of these *n*-alkyl lipids in the whole soil as well as across the grain size fractions. Moreover, it is also known that differently sized soil particles represent distinct microenvironments that contribute to the spatial heterogeneity and microbial diversity found in soils, where there is higher diversity and larger biomass of microbes in small size fractions than in coarse size fractions (Sessitsch et al., 2001; Hemkemeyer et al., 2018). This demonstrates that the microbial community is significantly affected by particle size and specific microbe-particle associations can greatly affect the distribution of lipid compounds in the soil.

It is reported in previous studies that the long-chain *n*-alkanes and *n*-alkanoic acids occur not only in leaves but also in other plant parts at varying concentrations (Jansen et al., 2007; Huang et al., 2011). To understand the fate of *n*-alkyl lipids compounds from soil with different vegetation cover, studies have mostly focused on the distribution patterns of individual lipid fractions like *n*-alkanes and *n*-alkanoic acids from the whole/bulk soil samples (Wiesenberg et al., 2004a; Wiesenberg et al., 2004b; Li et al., 2008). As different vegetation types are believed to control the contribution of OM in the soil, the distribution and preservation of the *n*-alkyl lipids across the grain size would not be similar under intrinsic soil forming conditions. Some other previous studies on particular vegetation cover showed information on *n*-alkyl lipids abundance and their distribution pattern across particle size fractions, which did not include the changes incurred by the variation in vegetation cover (trees or grasses). Previous studies, however, did not emphasize the fate of OM sources from different vegetation cover, leading to our limited understanding of the distribution pathways and preservation of *n*-alkyl

compounds in different ecosystems. Moreover, there is no possibility in comparison of lipid pools produced from different vegetation types under similar climatic conditions and similar soil textural attributes (Quenea et al., 2004; Clemente et al., 2011).

Partially modified plant material from the litter layer gets incorporated into the topsoil and early diagenetic modification of OM in the soil can induce significant changes in the original composition of these compounds. Combined with a link between vegetation cover quality and the microbial contribution to the stabilized OM across the grain size from different systems, several molecular indices derived from *n*-alkyl-compounds have large potential to exemplify the lipid sources, identification of preservation pathways and degradation patterns in the soil under different ecosystems. Similarly, these molecular indices can be used to ascribe the source which can be derived from plants and microorganisms and the degree of degradation of *n*-alkanes and *n*-alkanoic acids within the different size fractions. The two most common are carbon preference index (CPI) and average chain length (ACL). The CPI is a molecular ratio often used to assign the biological source and maturity of OM. High CPI values (>1) are used to indicate a terrestrial plant source (Eglinton and Hamilton, 1967). Average chain length of *n*-alkanes and *n*-alkanoic acids is the concentration weighted mean of carbon chain lengths to can be used to detect the shift in vegetation cover (Duan and He, 2011; Hughen et al., 2004; Zech et al., 2012). The predominance of odd over even (OEP) of *n*-alkanes and even over odd (EOP) of *n*-alkanoic acids can be used as proxy for lipid degradation pattern. High OEP and EOP values are characteristics of plant material, while the OEP/EOP values decrease with ongoing soil lipid modification (Buggle et al., 2010).

n-Alkanes and *n*-alkanoic acids associated in size fractions of soils under different vegetation cover represent valuable organic geochemical tools for investigation and reconstruction of past environments. Further, the persistence of *n*-alkyl lipids is important in sustaining the large OM pool in the context related to global climate change. Studying the degradation and preservation of these compounds in terms of molecular indices could increase the overall understanding of soil OM dynamics. Depending on the potential of molecular indices (CPI, ACL, OEP, EOP) to trace the OM incorporation from various sources and the residence time in soil, we can evaluate the applicability of *n*-alky lipids to trace the changes of carbon input and turnover rates across different size fractions which can provide valuable contribution in studies related to global climate change. Our concerns about soil OM dynamics have become an increasingly important consideration in the carbon cycle that can significantly have an impact on global warming.

Our present study from the biodiversity rich region of lower Gangetic Plain provides an opportunity to study the respective contribution of the various growth forms (forest, grassland and mixed system which consist of shrubs and grasses) to the soil lipids where the soil textural composition and climatic condition is similar. This is the first study describing the distribution of *n*-alkanes and *n*-alkanoic acids in particle size fractions, the association of these compounds with soil minerals particles and the opportunity to the application of molecular indices depending on the contribution of OM derived from plant and microbial source in soil size fractions from soil having different vegetation cover. Understanding the behavior these indices derive from *n*-alkanes and *n*-alkanoic acids in terms of their compartmentalization and preservation across the grain size of soil with different vegetation cover is useful to identify the incorporation pathways and the preservation mechanisms of OM which is of particular useful for paleoecological studies.

CHAPTER 2

Study area

The study area is an extensive fluvial setting from the region of Lower Gangetic floodplain of the Indian subcontinent, Nadia district (22.57°N and 87.31° E) West Bengal India (Roy et al., 2020). It is a highly biodiversity-rich area. The climatic conditions are warm and humid, having the mean annual temperature of 27 °C and annual rainfall of ~1200mm. The study area is marked by oxbow lakes and meander paleo-channels. The soils developed are moderately alkaline.

Lower Gangetic Plain is dominated by C_3 and C_4 plants that comprise of grasses and trees. These plants are major contributors to the soil OM. Plant waxes, a key photosynthetic product, provide key records of past environmental conditions because they can preserve the original molecular compositions that serve as promising proxy for past environmental reconstructions (Roy et al., 2020).

C_3 (trees) and C_4 (grass) plants

Atmospheric CO_2 enters through stomata and goes into mesophyll cells where CO_2 is fixed by an enzyme RuBisCo and produces phosphoglyceric acid, a three carbon compound reason for the name of the plants. Trees, shrub and cool season grasses are C_3 plants.

Similar to C_3 plants, atmospheric CO_2 enters through stomata and goes into mesophyll cells where it is fixed by an enzyme phosphoenolpyruvate (PEP) to make oxaloacetate, a 4 carbon atoms and hence the name C_4. CO_2 fixation in the outer layer of mesophyll cells to the inner layer of bundle sheath cells, which are able to concentrate CO_2 and thus fractionation of CO_2, less compared to C_3 plants. Warm season's grasses are C_4 plants.

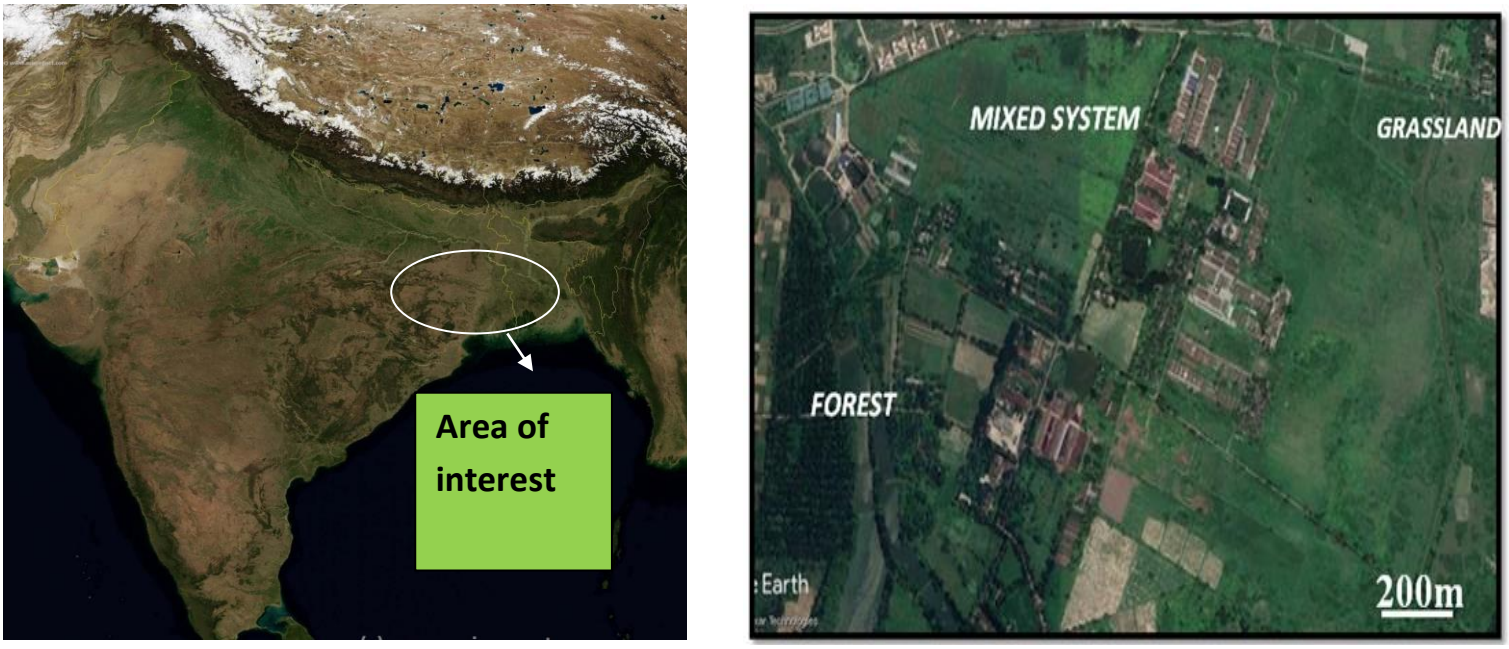


Figure 1: Left: Map of India showing the Lower Gangetic Plain; Right: Location of three different terrestrial systems from Lower Gangetic plain (Source: Google Earth)

CHAPTER 3

Materials and methods

3.1 Soil sampling

Dominant leaf samples from each sampling site of the forest, grassland and mixed systems were collected and preserved in ice-boxes before storing at -20°C. Before collecting the soil samples, litter layers were removed completely from the top soil surface. At each study site, the soil samples were collected (~400 g) after trenching a depth of 10-30 cm from five locations for an amalgamative representation of the forest, grassland and mixed systems. Nitrile gloves were used for sample collection were and the soil samples were collected in respective zip locks bags. After sampling, all material was stored in the freezer at -20°C and then the soil samples were freeze dried (Lyovapor L-200) and kept in sealed air-tight for further analysis. Aliquot of each sample (~30 g of soil) from individual sites were kept for grain size and clay mineralogy analysis. All roots were removed with tweezers and kept for further analysis.

Figure 2: Sampling site (a) Forest (b) Grassland (c) Mixed system

Table 1: Description of sampling location

Location description	Location 1	Location 2	Location 3
Location	Forest	Grassland	Mixed vegetation
	22°57.895′ N and 88°31.759′ E	22°57.783′ N and 88°31.580′ E	22° 57.189′ N and 88°31.167′ E
Precipitation(mm)	1200	1200	1200
Temperature (°C)	27	27	27
Lithology	Dark organic-rich clayey layer	Brownish yellow	Brownish yellow and reddish in color
Number of replicated plots	2	2	2
Soil properties			
Soil classification	Silty loam	Silty loam	Silty loam
Soil horizons	A	A	A
Soil depth (cm)	30 ± 3	32 ± 4	31 ± 5

3.2 Grain size analysis

Soil samples were treated 0.5N hydrochloric acid to remove carbonates. Subsequently, hydrogen peroxide was used to oxidize the organic matter. Hydrogen peroxide was added to the samples until all organic matter was oxidized and the reaction ceased. When the reaction ended, the samples were washed with distilled water by centrifugation at 3500 rpm for 5m. After that the supernatant was discarded. To measure the grain size distribution, treated soil samples from forest (n=4), grassland (n=6) and mixed system soil (n=4) were analyzed using Malvern Mastersizer 3000 laser-diffraction particle size analyzer at Indian Institute of Science Education and Research Education Kolkata, India.

3.3 Sieving analysis

Dry sieving method was applied to separate the soil (n= 6) into five different size fractions. The soil samples were passed through a series of stacked sieve meshes with defined opening sizes (ASTM). The soil is separated into fine sand (250-125µm), very fine sand (125-63 µm), coarse silt (63-32 µm), medium silt (32-20µm) and finer fraction (<20 µm).

3.4 X-ray Diffraction analysis (XRD)

To study the mineralogical compositions, all soil samples were first dispersed with sodium hexametaphosphate, $(NaPO_3)_6$ and kept for 7h to ensure the complete dispersion of sediment particles. The chemical dispersion also prevented grains from aggregating after sonication and during the grain-size measurements. The soil sample in solution was kept until the sediment settles out and only clay size particles are in suspension. The clay size fraction (< 2µm) in suspension was centrifuge at 5000 rpm for 5 min and the clay retrieved was prepared by oriented on glass slides. X-ray diffraction (XRD) were performed on samples and analyses were run on X-ray diffractometer, between 5 to 45°2Θ using a step size of 0.02°2Θ and a count time of at least 4s per step. Clay minerals were characterized by their interlayer spacing as revealed by XRD analysis.

3.5. Scanning Electron Microscopy (SEM) analysis

The characterization of microaggregates at the micrometer scale using Scanning Electron Microscopy (SEM) provides useful information on the association of OM with the primary particles. Sand-size (>125µm) and finer fractions (<20µm) were analyzed by using the SEM facility at the Indian Institute of Science Education and Research (IISER) Kolkata. The images were taken through SUPRA 55 VP-41-32 instrument with Smart SEM Zeiss software.

3.6 Quantification and characterization of *n*-alkanes and *n*-alkanoic acids

The total lipids from the individual whole soil of three different systems and five particle size fractions (~10 g) were extracted for two cycles using DIONEX accelerated solvent extractor (ASE-350) with dichloromethane and methanol (93:7) at temperature of 100° C and pressure of 1600 pound square inch. The extracted lipid is concentrated using the rotary evaporator. *n*-Alkanes was separated from the total extracted lipid using short column chromatography filled with silica gel in a glass pipette plugged by

quartz wool and hexane was used as an eluent. The aliquots remains were eluted with dichloromethane and methanol (2:1) for *n*-alkanoic acid extraction.

Subsequently, the alkanoic acids portion was saponified with 1 N KOH in methanol by heating in a water bath at 70 °C for 2h. After addition of sodium chloride (NaCl) and DCM treated water, a neutral fraction was extracted by hexane. The pH is then lowered to <2 by addition of HCl and the acid fraction were extracted with hexane. BF_3-methanol was added to the extracted solution to convert the mixture to fatty acid methyl esters (FAME) by heating at 70 °C for 45 min. FAME eluted with hexane was taken out using a glass pipette and passed through Na_2SO_4 crystals to absorb the moisture. Compound identification and quantification of *n*-alkanes and *n*-alkanoic acids was performed via GC-MS analysis.

Gas chromatography (GC Aligent 7890 A) equipped with split/split less injector and non-polar column (HP-5, 30m × 30mm × 0.25micro m) attached with the mass spectrometer (MS). The samples were injected in 1:1 split mode with oven temperatures that was increased from 60°C to 320°C at 8°C per min with 12 min hold time using the helium gas as carrier (1 µl/min). The reference standards, (SUPELCO C_8 to C_{40} Alkanes and Indiana A5) and five fatty acids methyl esters (FAME) with known concentrations were used. Identification of peak was based on retention times and mass spectra using the NIST library.

3.7 *n*-Alkane and *n*-alkanoic acid indices

Total *n*-alkanes and *n*-alkanoic acids concentrations were calculated as the sum of C_{11} to C_{33} and C_{12} to C_{34} (odd as well as even ones), respectively, and given in ng of lipids/g of dry sediments. High molecular weight (HMW_{alk}) and Low molecular weight (LMW_{alk}) is the sum of the abundance of *n*-alkanes with C_{23} to C_{33} and C_{11} to C_{23} (odd as well as even ones) respectively. High molecular weight (HMW_{acid}) and Low molecular weight (LMW_{acid}) is the sum of the abundance of *n*-alkanoic acids with C_{22} to C_{34} and C_{12} to C_{21} (odd as well as even ones) respectively.

Total *n*-alkane concentration =$\sum C_{11} - C_{33}$

Total *n*-alkanoic acid conc=$\sum C_{12} - C_{34}$

Low Molecular Weight (LMW_{alkane}) = $\sum C_{11} - C_{22}$

High Molecular Weight (HMW_{alkane}) =$\sum C_{23} - C_{33}$

Low Molecular Weight ($LMW_{alkanoic\text{-}acid}$) =$\sum C_{12} - C_{21}$

High Molecular Weight ($HMW_{alkanoic\text{-}acid}$) = $\sum C_{22} - C_{34}$

Carbon Preference Index for long and short-chain *n*-alkanes ($CPI_{alk(23-33)}$ and $CPI_{alk(15-23)}$) is the measures of the relative abundance of odd-over-even carbon chain lengths. Whereas, carbon preference index for long and short chain *n*-alkanoic acids ($CPI_{acid(22-32)}$ and $CPI_{acid(12-20)}$) is the measures of the relative abundance of even-over-odd carbon chain lengths (Marzi et al., 1993).

$$CPI_{alk\ (23-33)} = \frac{((C_{23}+C_{25}+C_{27}+C_{29}+C_{31})+ (C_{25}+C_{27}+C_{29}+C_{31}+C_{33}))}{(2*(C_{24}+C_{26}+C_{28}+C_{30}+C_{32}))} \quad \text{.....Equation 1}$$

$$CPI_{alk\ (15-23)} = \frac{(C_{15}+C_{17}+C_{19}+C_{21})+(C_{17}+C_{19}+C_{21}+C_{23})}{(2*(C_{16}+C_{18}+C_{20}+C_{22}))} \quad \text{..... Equation 2}$$

$$CPI_{acid\ (22-32)} = \frac{(C_{22}+C_{24}+C_{26}+C_{28}+C_{30}) + (C_{24}+C_{26}+C_{28}+C_{30}+C_{32})}{(2*(C_{23}+C_{25}+C_{27}+C_{29}+C_{31}))} \quad \text{.....Equation 3}$$

$$CPI_{acid\ (12-20)} = \frac{(C_{12}+C_{14}+C_{16}+C_{18}) + (C_{14}+C_{16}+C_{18}+C_{20})}{(2*(C_{13}+C_{15}+C_{17}+C_{19}))} \quad \text{..... Equation 4}$$

The average chain length (ACL) of *n*-alkyl was determined by modifying the equation of (Poynter et al., 1989). We used the odd chain lengths (*n*-C_{25} to *n*-C_{33}) for the alkanes and even chain lengths (*n*-C_{24} to *n*-C_{32}) for the alkanoic acids.

$$ACL_{alk\ (25-33)} = \frac{(25*C_{25}+27*C_{27}+29*C_{29}+31*C_{31}+33*C_{33})}{(C_{25}+C_{27}+C_{29}+C_{31}+C_{33})} \quad \text{.....Equation 5}$$

$$ACL_{acid\ (24-32)} = \frac{(24*C_{24}+26*C_{26}+28*C_{28}+30*C_{30}+32*C_{32})}{(C_{24}+C_{26}+C_{28}+C_{30}+C_{32})} \quad \text{.....Equation 6}$$

The odd-over-even predominance (OEP) of the long chain *n*-alkanes and the even-over-odd predominance (EOP) for the long-chain *n*-alkanoic acids can be used as a proxy to identify the pattern of degradation (Li et al., 2018).

$$OEP = \frac{(C_{27}+C_{29}+C_{31}+C_{33})}{(C_{26}+C_{28}+C_{30}+C_{32})} \quad \text{..... Equation 7}$$

$$EOP = \frac{(C_{20}+C_{22}+C_{24}+C_{26}+C_{28}+C_{30})}{(C_{19}+C_{21}+C_{23}+C_{25}+C_{27}+C_{29})} \quad \text{.....Equation 8}$$

CHAPTER 4

Results

4.1 Grain-size and mineralogical distribution

Grain-size distribution analysis measured from laser diffraction was recovered for most of the bulk samples (> 95%) for all soil samples (Table 2). The majority of soil occurred in the 63-32μm sizes (silt fraction); accounting for ca. 60%, 67% and 62% of the whole soil sample from forest, grassland and mixed systems soil respectively. Together in all the systems silt dominated the sediment mass. Of the residual, ca. 30% resided in sand-size and ca. 3%, in the clay-size fractions.

The soils have mineralogical composition dominated by quartz, feldspar and the clay minerals assemblages mainly composed of illite and kaolinite. Illite showed a basal peak at 7Å and 17 Å. Kaolinite is characterized by peaks at 13 Å and 25 Å, quartz and feldspar are characterized by 28 Å. The mineralogical distribution pattern in all the three systems is similar (Figure. 3).

Table 2: Volume contents of sand, silt, and clay of GSD at different sampling sites in our study

Soil Sample	% Sand±SE	% Silt±SE	% Clay±SE
Forest	37.3 ± 5.23	60.7 ± 4.37	2.0 ± 1.13
Grassland	28.9 ± 3.27	67.8 ± 7.34	3.2 ± 0.57
Mixed vegetation	35.8 ± 11.18	62.2 ± 10.33	1.9 ± 0.87

SE=Standard error; (n=7)
GSD =Grain size distribution

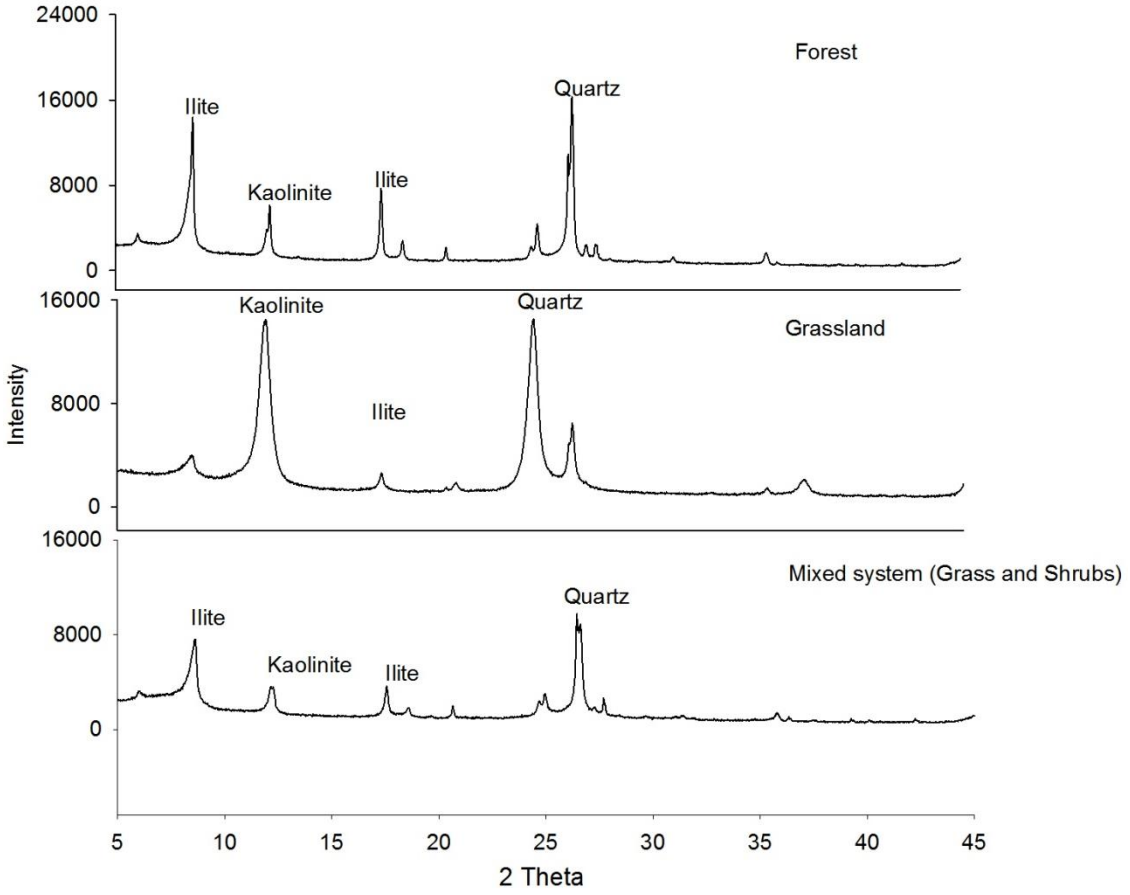


Figure 3: Graphical output showing the measured XRD pattern of clay minerals in the forest (top), grassland (middle) and mixed system (bottom) soils

4.2 Particle size separation

Together, the five different size fractions of soil separated from forest, grassland and mixed systems they represent ~98% of the whole soil, indicating a high recovery. In agreement with the silty nature of these soils from all the systems, physical particle size fractionation showed that the silt fraction (63-20µm) from forest, grassland and mixed systems accounts ca. 67 wt% from the initial weight of whole soil, whereas the weight of the sand-size (250-63µm) and finer fractions (<20µm) reached ca. 25 wt% and ca. 8 wt% (Table 3).

Straight chain *n*-alkyl lipids from whole soils and particle-size separates soils were normalized to soil mass (ng of *n*-alkyl lipids/g dry soils). When the absolute concentration (ng/g) of lipids is considered, it was found that silt-size fractions (63-20µm) are very poor (~14 wt %) in *n*-alkyl lipids contents relative to whole soil lipid content (Figure 4; Table 3), whereas the sand-size (250-63µm) are comparatively

higher in amount (~40 wt %) (Figure 4; Table 3). A substantially higher value is noticed for the finer fractions of soil (~90 wt %) compared to the sands and silts (Figure 4; Table 3). Normalized data of the particle size separates revealed that the sum of the *n*-alkyl content in the particle size was larger than the *n*-alkyl lipid content of the respective whole soils (Figure 4).

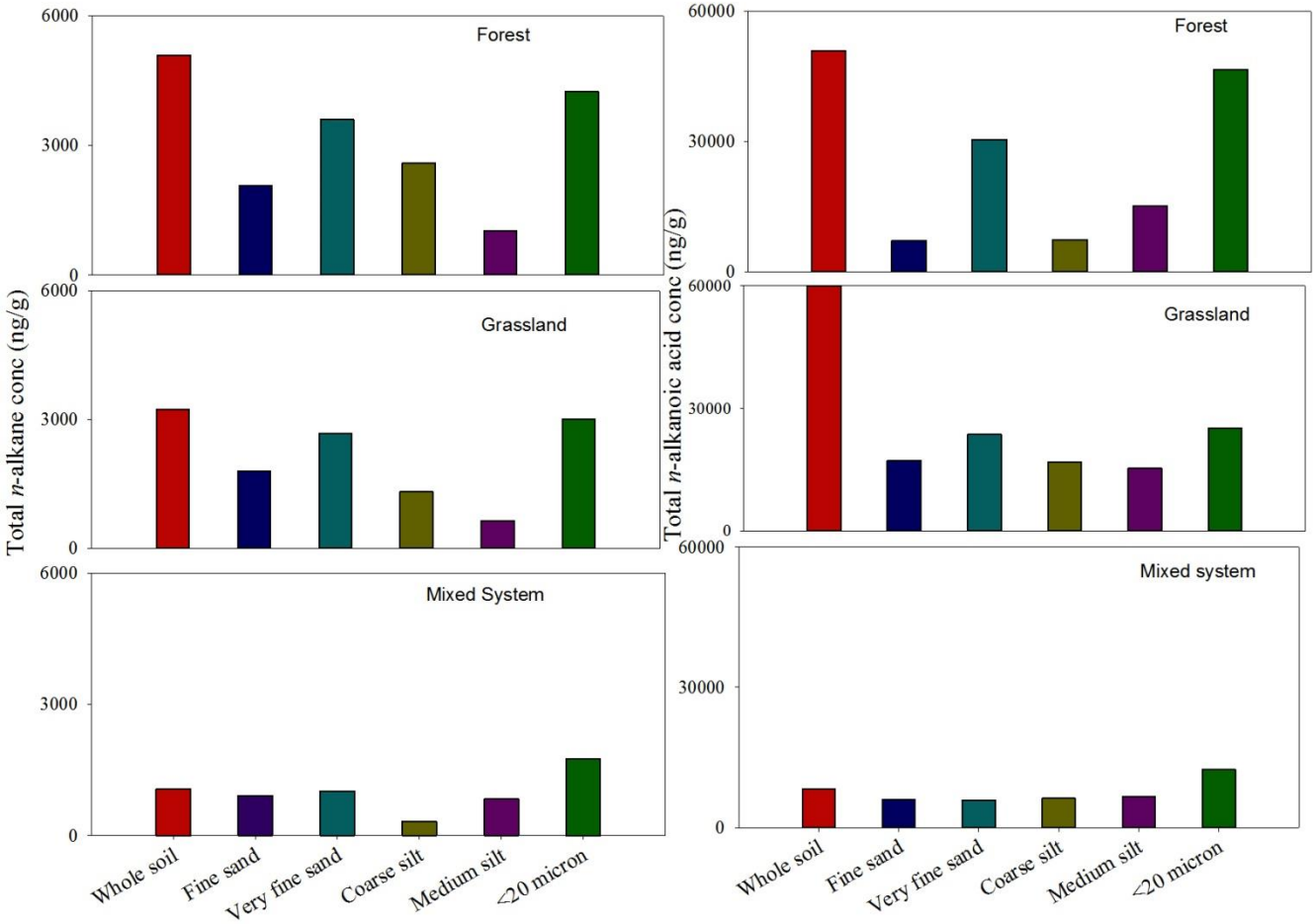


Figure 4: Total concentrations of *n*-alkanes and *n*-alkanoic acids in whole soils of forest, grassland and mixed system and in different size fractions

Table 3: ***n*-Alkyl lipids content of whole soil and size fractions**

Soil	Size separate/Micron	Mass ± SE/% Bulk	*n*-alkane extract yield (ng/g)	*n*-alkanoic extract yield (ng/g)
Fine sand (250-125µm)	Very fine sand (125-63µm)	Coarse silt (63-32µm)	Medium silt (32-20µm)	Finer fraction (< 20µm)
Forest	Fine sand	10 ± 9	2079.7	7114.9
	Very fine sand	18 ± 8	3606.0	30475.3
	Coarse silt	46 ± 7	2601.2	7384.6
	Medium silt	20 ± 5.2	1027.7	15165.5
	< 20µm	5 ± 1	4253.2	46492.0
	% recovery	99%		
	Bulk yield (ng/g)		5090.9	50964.3
Grassland	Fine sand	11 ± 6	1793.7	17259.4
	Very fine sand	18 ± 4	2669.0	23550
	Coarse silt	46 ± 3.1	1321.4	16884.6
	Medium silt	15 ± 5	647.8	15285.6
	< 20µm	9± 1	3008.7	25068.8
	% recovery	98%		
	Bulk yield (ng/g)		3234.7	69068.6
Mixed system	Fine sand	12 ± 5	913.3	6017.3
	Very fine sand	17 ± 6.5	1015.3	5871.6
	Coarse silt	40 ± 7	326.7	6210.5
	Medium silt	14 ± 3.1	835.5	6599.2
	< 20µm	8± 1.5	1749.7	12396.6
	% recovery	96%		
	Bulk yield (ng/g)		1063.6	8270.5

SE= Standard error; Number of analysis varied from 2 to 3
ng/g= nanogram of lipid/gram of dry soil

4.3 *n*-Alkanes abundances and distribution patterns of whole soils and particle size fractions

Differences in molecular abundance and distribution patterns of *n*-alkanes observed in whole soil between forest, grassland and mixed systems and size fractions are illustrated in Figure 5. Comparison between the three systems, the total *n*-alkane concentration (C_{11}-C_{33}) is significantly higher in forest (5.09μg/g) than in grassland (3.23μg/g) and mixed systems (1.06μg/g). Forest soil have the highest higher molecular weight ($HMW_{alk(23-33)}$) *n*-alkanes compounds of 4.5μg/g than grassland (2.5μg/g) and mixed system (0.7μg/g), whereas low molecular weight ($LMW_{alk(11-22)}$) *n*-alkanes is larger in grassland (0.6μg/g) than in forest (0.5μg/g) and mixed system (0.3μg/g) (Table 1a; Table 1a; Table 1c; Figure 5).

Across the five different size fractions, *n*-alkanes concentration varied in a similar way in all the three systems, lowest for the silt-size fraction and the highest for the finer fraction (Table 1a; Table 1b; Table 1c; Figure 5). Total *n*-alkanes concentration in the forest ranges from 1.02μg/g to 4.35μg/g (Table 1a). Relatively low total *n*-alkane concentrations are observed in coarse silt (63-32μm) and medium silt (32-20μm) (1.02μg/g) (Table 1a; Figure 5). Fine sand (250-125μm) and very fine sand (125-63μm) (3.60μg/g) have comparatively higher concentration of *n*-alkanes compared to silt-size but lower than that of finer fractions (<20μm) (4.35μg/g) (Table 1a; Figure 5). Higher total *n*-alkanes concentration in sand fractions mostly contributed from the $HMW_{alk(23-33)}$ compounds. In grassland, the total *n*-alkanes concentration in size fractions varied between 0.64μg/g to 3.00μg/g (Table 1b). Similarly, in this system, coarse silt (63-32μm) and medium silt (32-20μm) (0.64μg/g) have the lowest *n*-alkanes concentration compared to sand (250-63μm) (2.66μg/g) and finer fractions (<20μm) (3.00μg/g) (Table 1b; Figure 5). The total *n*-alkanes concentration in mixed system ranges from 0.32μg/g to 1.74μg/g (Table 1c). Lowest concentration for the coarse silt (63-32μm) and medium silt (32-20μm) (0.32μg/g) size fractions and the largest for the finer fractions (<20μm) (1.74μg/g) are observed (Table 1c; Figure5).

Finer fractions have higher amounts of $HMW_{alk(23-33)}$ and $LMW_{alk(11-22)}$ compounds than sand- and silt-size fractions in all the systems. Another striking observation is the relatively high concentration of $HMW_{alk(23-33)}$ in sand fractions of 3.2μg/g, 1.9μg/g and 0.6μg/g than silt-size fractions each with an amount of 2.1μg/g, 0.9μg/g and 0.1μg/g in forest, grassland and mixed system respectively. Silt-size fractions, however, have the lowest total *n*-alkanes concentration and lowest $HMW_{alk(23-33)}$ content in all the systems (Table 1a; Table 1b; Table 1c; Figure 6).

A homologous series of *n*-alkanes (C_{11}-C_{33}) of whole soils and five particle size fractions from forest, grassland and mixed systems are identified (Table 1a; Table 1b; Table 1c; Figure 5). *n*-Alkanes distribution patterns of the whole soil from all the three systems shows predominance of odd-over-even carbon number in the long-chain, which is the characteristics of higher terrestrial plants. Within short-chains ($<C_{23}$), which can be derived from modification of long-chain to short-chain compounds and contribution from the microbial communities, a preference of odd/even homologues is missing. Soils from all the systems show an abundance of odd carbon number long-chain C_{27}, C_{29}, C_{31} and C_{33} *n*-alkanes. Forest soils have a maximum carbon number at C_{29} *n*-alkanes. C_{33} dominated in grassland and mixed systems. The representative *n*-alkanes spectra and various chain lengths of the whole soil from forests, grasslands mixed systems are illustrated in Figure 4a.

The *n*-alkanes distribution patterns of particle size fractions revealed the same as that of their respective whole soils (Figure 5). Sand-, silt-size and finer fractions separated from forest, grassland and mixed systems soil have higher abundance of long-chain compounds (sum of both even and odd) than that of short-chains compounds (Table 1a; Table 1b; Table 1c; Figure 5). For all the size fractions from sand-size to finer fractions, C_{31} *n*-alkane is the most abundant homologue for forest, grassland and mixed systems soil. The abundance of other odd carbon number long-chain *n*-alkanes is lower than that of the C_{31}. A marked difference was observed in the distribution patterns of *n*-alkanes between the sand-, silt-size and the finer fraction of the forest, grassland and mixed systems. Sand fractions exhibit a significant odd to even carbon number predominance in the long-chain compared to silt-size and finer fractions. Notably, no odd-over-even carbon number predominance in the short-chain (C_{11}-C_{21}) of the sand fractions. Long chain *n*-alkanes from silt and finer fractions of forest, grassland and mixed system show odd-over-even carbon number predominance but not as strong as in the sand fractions, whereas short-chains lack clear odd to even carbon number predominance. The *n*-alkanes of finer fractions of soil show a bimodal distribution with a maximum at C_{16} and a sub-maximum at C_{14}/C_{18} in the short-chain and C_{31} in the long-chain. Again, no predominance of odd/even noted for the short-chain *n*-alkanes in finer fractions. The long-chain is predominantly odd carbon numbers in the C_{23}-C_{33} range. The representative *n*-alkanes spectra and various chain lengths of the sand-, silt- and finer size-fractions are given in Figure 4b.

4.3.1 Molecular proxies of *n*-alkanes

Several molecular parameters were derived for *n*-alkanes to enable the identification of incorporation and preservation pathways of these compounds in whole soil and particle size fractions. The CPI was calculated for long-chain ($CPI_{alk(23-33)}$) and for short-chain ($CPI_{alk(15-21)}$) *n*-alkanes separately, using the equation 1 and 2 respectively. The CPI values of both indicate the preference of odd carbon numbers *n*-alkanes over the even counterparts. The long chain *n*-alkanes $CPI_{alk(23-33)}$ values are 7.0, 5.2 and 2.1 in forest, grassland and mixed systems whole soil respectively (Table 1a; Table 1b; Table 1c Figure 6).

Across the five different particle size fractions, forest, grassland and mixed systems show a decreasing trend of 30-60% in $CPI_{alk(23-33)}$ (Table 1a; Figure 6) with decreasing grain size from the sand to the finer fractions. $CPI_{alk(23-33)}$ values of long-chain *n*-alkanes varied between 7 and 10 for sand-size fractions which is comparatively higher than $CPI_{alk(23-33)}$ values with range of 3 and 4 in the finer fractions. $CPI_{alk(23-33)}$ for silt-size fractions lies between sand and finer fractions (Table 1a; Table 1b; Table 1c Figure 6).

The short-chain *n*-alkanes $CPI_{alk(15-21)}$ values are 0.5, 0.4 and 0.4 in forest, grassland and mixed system whole soils respectively. Short-chain *n*-alkanes $CPI_{alk(15-21)}$ varied between 0.1-0.5 across the size fractions of all systems. Lowest $CPI_{alk(15-21)}$ values are observed in silt-size and finer fractions and relatively higher values in the sand-size (Table 1a; Table 1b; Table 1c; Figure 6).

Average chain length ($ACL_{alk(25-33)}$) was calculated for long-chain *n*-alkanes using the equation 3. $ACL_{alk(25-33)}$ values of whole soil are 29.8, 30.6 and 29.4 in forest, grassland and mixed system respectively. Slight decreasing patterns in $ACL_{alk(25-33)}$ values of ~1-2% from sand to finer fractions in all systems are identified. Forest soils have $ACL_{alk(25-33)}$ values of 30.2 in sand size and 29.7 in finer fractions. Grasslands and mixed systems have $ACL_{alk(25-33)}$ values of 30.3 and 29.6 in sand and 27.9 and 29.3 in finer fractions respectively (Table 1a; Table 1b; Table 1c; Figure 6).

C_{29}/C_{31} was calculated by dividing the *n*-alkanes concentration of C_{29} by C_{31}. Forest, grassland and mixed systems soil have C_{29}/C_{31} of 1.3, 0.8 and 0.8 in respectively (Table 1a; Table 1b; Table 1c). C_{29}/C_{31} shows a negative correlation with size fractions of all systems. Forest soil has C_{29}/C_{31} of 0.5 in sand and 0.7 in finer fractions. C_{29}/C_{31} ratio increases from 0.3 to 0.5 and 0.9 to 1 from sand-size to finer fractions in grassland and mixed system respectively.

Odd-over-Even Predominance ($OEP_{(27-33)}$) is the measure of the relative abundance of odd to even carbon numbers long-chains by using the concentrations of a given *n*-alkanes chain length. $OEP_{(27-33)}$ was calculated for long-chain *n*-alkane whole soil and size fractions from the equation 4. Lower $OEP_{(27-33)}$ of 6.0 and 2.5 is identified in grassland and mixed system whole soil respectively than that in forest soil with value of 7.3. $OEP_{(27-33)}$ shows a decrease of ~40-50% (9.0 to 4.1), (9.0 to 3.8) and (3.5 to 2.3)

from sand-size to finer fractions of forest, grassland and mixed system soil respectively (Table 1a; Table 1c; Figure 6). $OEP_{(27\text{-}33)}$ appears to resemble the decreasing trend as $CPI_{alk(23\text{-}33)}$.

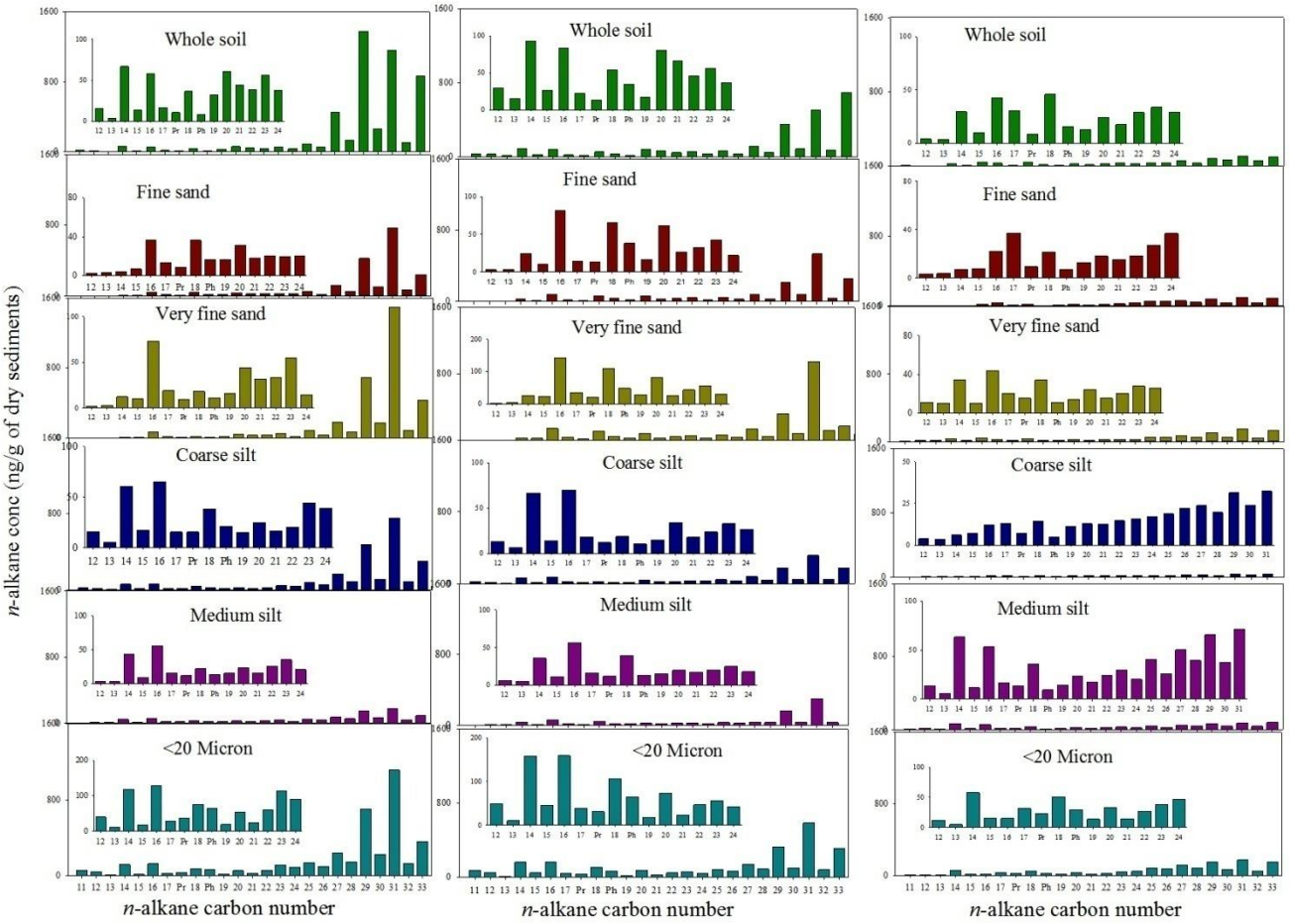


Figure 5: Abundances of the C_{11}–C_{33} *n*-alkanes (in ng *n*-alkanes/g dry soil) and the resulting *n*-alkanes distribution patterns in the forest, grassland and mixed system soils (left, middle and right-hand side, respectively). The inset image illustrates the results for the *n*-alkanes chain lengths (*n*-C_{11}–C_{24}) at a scale adapted to their low concentrations.

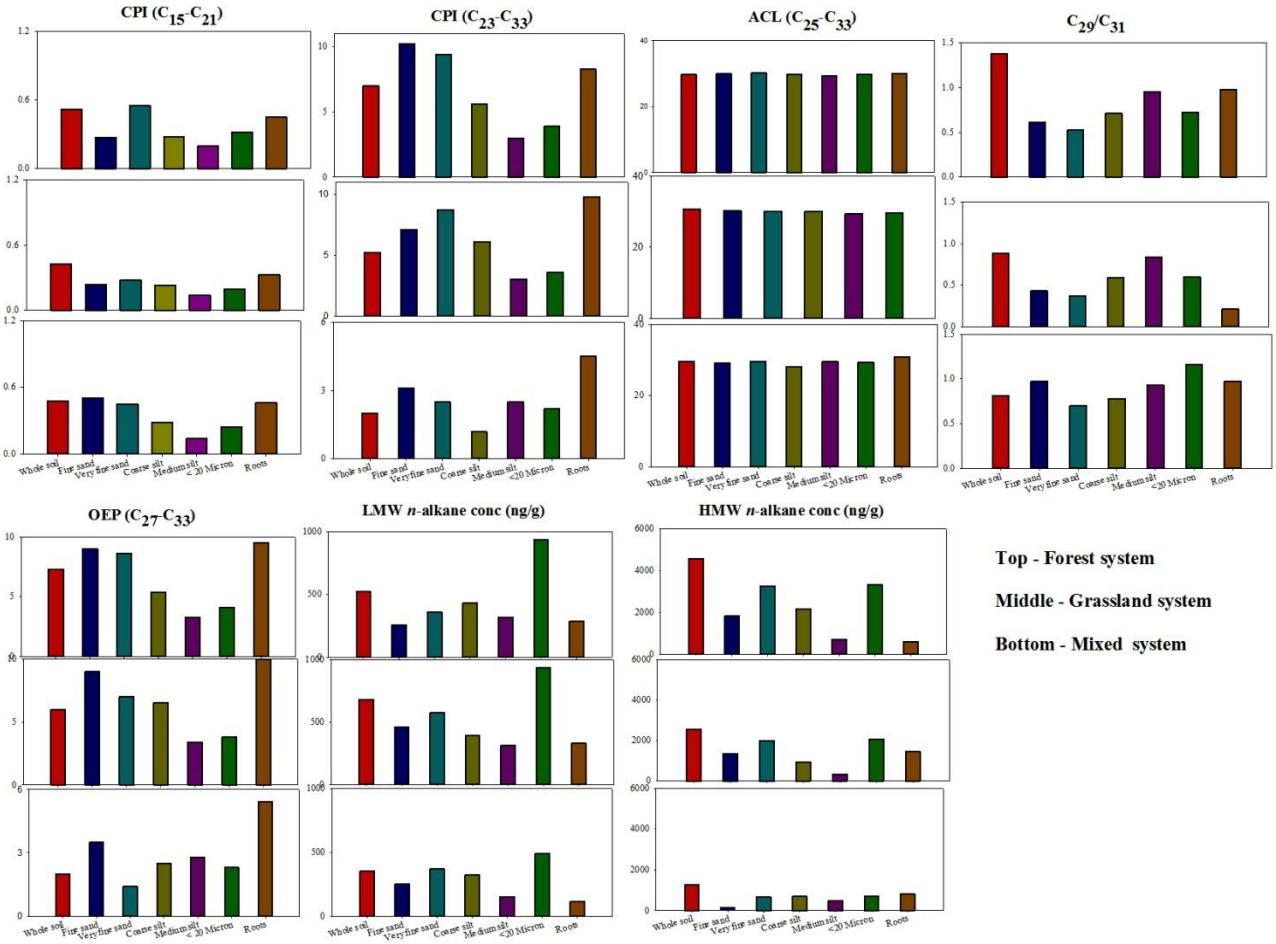


Figure 6: *n*-Alkanes molecular indices and the relative proportions of the short and long-chain *n*-alkanes for the investigated soil systems (Top – Forest system, Middle – Grassland and Bottom – Mixed system)

Table 1 a: Concentrations and other parameters of *n*-alkanes compounds in forest soil and roots

Soil	Total alk conc	$CPI_{(23-33)}$	$CPI_{(15-21)}$	$ACL_{(25-33)}$	OEP	C_{29}/C_{31}	Cmax	LMW	HMW	LMW/HMW
Whole soil	5090.9	7.0	0.5	29.8	7.3	1.3	C29	521.1	4569.8	0.1
Fine sand	2079.7	10.2	0.2	30.1	9.0	0.6	C31	256.0	1823.7	0.1
Very fine sand	3606.0	9.4	0.5	30.2	8.6	0.5	C31	356.0	3249.9	0.1
Coarse silt	2601.2	5.6	0.2	29.9	5.4	0.7	C31	430.8	2170.3	0.1
Medium silt	1027.7	3.3	0.1	29.3	3.3	0.7	C31	316.8	710.9	0.4
<20 Micron	4253.2	3.9	0.3	29.7	4.1	0.9	C31	931.9	3321.3	0.2
Roots	897.2	8.2	0.9	30	9.5	0.9	C31	283.4	613.7	0.4

Table 1 b: Concentrations and other parameters of *n*-alkanes compounds in grassland soil and roots

Soil	Total alk conc	$CPI_{(23-33)}$	$CPI_{(15-21)}$	$ACL_{(25-33)}$	OEP	C29/C31	Cmax	LMW	HMW	LMW/HMW
Whole soil	3234.7	5.2	0.4	30.6	6.0	0.8	C33	680.8	2553.8	0.2
Fine sand	1793.7	7.1	0.2	30.3	9.0	0.4	C31	458.5	1335.2	0.3
Very fine sand	2669.0	8.7	0.2	30.1	7.0	0.3	C31	673.5	1995.4	0.3
Coarse silt	1321.4	6.1	0.2	30.0	6.5	0.5	C31	393	928.4	0.4
Medium silt	647.8	3.0	0.1	29.3	3.4	0.5	C31	312	335.7	0.9
<20 Micron	3008.7	3.6	0.1	29.7	3.8	0.8	C31	936.1	2072.5	0.4
Roots	1770.5	9.7	0.3	32.0	13.7	0.2	C33	327.3	1443.2	0.2

Table 1 c: Concentrations and other parameters of *n*-alkanes compounds in mixed system soil and roots

Soil	Total alk conc	$CPI_{(23-33)}$	$CPI_{(15-21)}$	$ACL_{(25-33)}$	OEP	C_{29}/C_{31}	Cmax	LMW	HMW	LMW/HMW
Whole soil	1063.6	2.1	0.4	29.4	2.5	0.8	C31	355.9	707.6	0.5
Fine sand	913.3	3.1	0.1	29.2	3.5	0.9	C31	255.7	657.5	0.3
Very fine sand	1015.3	2.5	0.2	29.6	1.4	0.7	C31	326.5	688.7	0.4
Coarse silt	326.7	1.2	0.4	28.1	2.5	0.7	C31	155.2	171.4	0.9
Medium silt	835.5	2.5	0.1	29.5	2.8	0.9	C31	370.8	464.7	0.7
<20 Micron	1749.7	2.2	0.2	29.3	2.3	1.1	C33	491.1	1258.6	0.3
Roots	928.9	4.5	0.4	30	5.4	0.4	C33	118.0	810.7	0.1

4.4 *n*-Alkanoic acid abundance and distribution patterns of whole soils and particle size fractions

n-Alkanoic acid molecular abundances and distribution patterns in soil under different vegetation cover are illustrated in (Table 2a; Table 2b; Table 2c; Figure 7). A significantly higher total *n*-alkanoic acid concentration (C_{12}-C_{34}) of 69.06μg/g is found in grassland whole soil than forest (50.96μg/g), and mixed system (8.27μg/g). However, $HMW_{acid(22-34)}$ is highest in forest (29.9μg/g) intermediate in grassland (23.9μg/g) and lowest in mixed system soil (2.4μg/g). Higher $LMW_{acid\ (12-21)}$ content of 45.1μg/g in grassland than forest and mixed systems of 21.0μg/g and 5.8μg/g respectively is observed (Table 2a; Table 2b; Table 2c; Figure 8). Representative *n*-alkanoic acid spectra and various chain lengths of the whole soil between forests, grasslands mixed systems are illustrated in Figure 4c.

The total *n*-alkanoic acid concentration in size fractions separated from forest soil ranges from 7.11μg/g, 30.47μg/g, 7.38μg/g, 15.16μg/g and 46.49μg/g in fine sand, very fine sand, coarse silt, medium silt and fine fractions respectively (Table 2a; Figure 7). Grassland varied between 17.25μg/g, 23.55μg/g, 16.88μg/g, 15.28 μg/g and 25.06μg/g in fine sand, very fine sand, coarse silt, medium silt and finer fractions respectively (Table 2b; Figure 7). The total *n*-alkanoic acid concentration of mixed system size fractions ranges from 6.01μg/g, 5.87μg/g, 6.21μg/g, 6.59μg/g and 12.39μg/g in fine sand, very fine sand, coarse silt, medium silt and fine fractions respectively (Table 2c; Figure 6).

The total *n*-alkanoic acid, $HMW_{acid(22-34)}$ and $LMW_{acid(12-21)}$ compounds are most abundant in the finer fractions of all the systems (Table 2a; Table 2b; Table 2c). Sand-size fractions in forest of 15.1μg/g, grassland of 2.4μg/g and mixed system of 1.5μg/g accounted a relatively larger amount of $HMW_{acid(22-34)}$ compared to silt fractions of 2.0μg/g, 1.6μg/g and 0.5μg/g in forest, grassland and mixed systems respectively. When compared with the finer fractions, $HMW_{acid(22-34)}$ of sand-size is much lower in abundance.

Distribution patterns of *n*-alkanoic acid (C_{12}-C_{34}) in the forest, grassland and mixed systems are characterized by strong even over odd carbon numbers predominance, with maximum carbon number at C_{16}, followed by C_{18}. Patterns of bimodal distribution with maxima at C_{16} in the short-chain and C_{24} in the long-chain of whole soils from all systems are identified. A relatively larger amounts of ubiquitous C_{16} and C_{18} each in an amount of 19.5µg/g and 9.5µg/g respectively is observed in grassland whole soil than in forest (8.6µg/g and 3.2µg/g) and mixed system (3.8µg/g and 0.7µg/g) (Table 2a; Table 2b; Table 2c; Figure 7).

Forest, grassland and mixed system particle size fractions show the distribution patterns of *n*-alkanoic acids that are similar to the patterns in their respective whole soils. The distributions are characterized by even over odd carbon number predominance *n*-alkanoic acids (Figure 7). Across the different size

fractions of all the three systems, *n*-alkanoic acids distributions are dominated by the short-chain compounds (<C_{20}), but the long-chain (>C_{22}) compounds are also significant. The patterns of bimodal distributions with maxima at C_{16} in the short-chain and C_{24}/C_{26} in the long-chain across the size fractions are identified. A significant even-over-odd carbon number in the long-chain and short-chain *n*-alkanoic acids are observed for sand-, silt-size and finer fractions. Overall, the distribution of long-chain *n*-alkanoic acids in whole soils and in size fractions shows a decreasing abundance from C_{24} to C_{34}. The decreasing rate of C_{24} to C_{34} is higher in the silt and finer fractions compared to the coarse fractions of all the systems (Figure 7). Representative *n*-alkanoic acid spectra and various chain-lengths of the sand-, silt- and finer size-fractions are given in Figure 4d.

4.4.1 Molecular proxies of *n*-alkanoic acids

The carbon preference index was calculated for the long-chain ($CPI_{acid(22-32)}$) and short-chain ($CPI_{acid(12-20)}$) *n*-alkanoic acids using the equation 5 and 6 respectively, to identify the possible behavior of sources in the short and long-chain compounds. $CPI_{acid(22-32)}$ values of long-chain *n*-alkanoic acids are 2.4, 2.6 and 3 in forest, grassland and mixed system whole soils respectively.

In size fractions, $CPI_{acid(22-32)}$ in forest and mixed systems shows a decreasing trend of ~15-20% with decreasing size fractions (Table 2a; Table 2c; Figure 8), whereas grassland show a decreasing trend of only ~5% from sand to finer fractions (Table 2b; Figure 8). The % of decrease in $CPI_{acid(22-32)}$ of 4 to 2.4 and 4.7 to 3.5 from sand to finer fractions in forest(~25%) and mixed system (~35%) respectively is higher when compared to grassland (~15%) where the decrease in $CPI_{acid(22-32)}$ is only from 4.5 to 4.3 (Figure 8).

Short-chain $CPI_{acid(12-20)}$ values of 10.4, 19.9 and 20.4 in forest, grassland and mixed system soils respectively are higher than their respective long-chain $CPI_{acid(22-32)}$ (Figure 8). $CPI_{acid(12-20)}$ values in five different size fractions in all the three systems are ~60-70% higher than the long chain $CPI_{acid(22-32)}$ values (Table 2a; Table 2b; Table 2c; Figure 8). $CPI_{acid(12-20)}$ value is much higher in all the five size fractions than that in their respective whole soils, mostly in the grassland and mixed system (Table 2b; Table 2c; Figure 8). Higher $CPI_{acid(12-20)}$ in sand-size fractions of 31-37 for grasslands and 53-47 for mixed vegetation soils and lower $CPI_{acid(12-20)}$ of 26.1 and 16.8 in finer fractions for grasslands and mixed system soils are identified. The $CPI_{acid(12-20)}$ values in the range of 10-13 for forest soil behave similarly across grain size (Figure 8).

Average chain length ($ACL_{acid(24-32)}$) was calculated using the equation 8 for long-chain *n*-alkanoic acids. Forest, grassland and mixed system whole soil have $ACL_{acid(24-32)}$ values of 27.7, 27.5 and 27.2

respectively. Across five different size fractions, $ACL_{acid(24-32)}$ values show a slight (~3%), but non-significant decreasing trend with decreasing grain size in all the systems (Figure 8).

Even over odd predominance ($EOP_{(22-34)}$) was calculated for long-chain *n*-alkanoic acids from equation 7. $EOP_{(22-34)}$ of 3.5, 3.6 and 3.9 are identified in forest, grassland and mixed system whole soil respectively. A decrease of ~20-30% in $EOP_{(22-34)}$ from sand-size to finer fractions in all the soil systems are observed (Figure 8). Forest, grassland and mixed systems have $EOP_{(22-34)}$ of 4.5, 4.8 and 5.5 in sand-size and 3.5, 3.6 and 3.4 in finer fractions respectively (Table 2a, 2b and 2c; Figure 8).

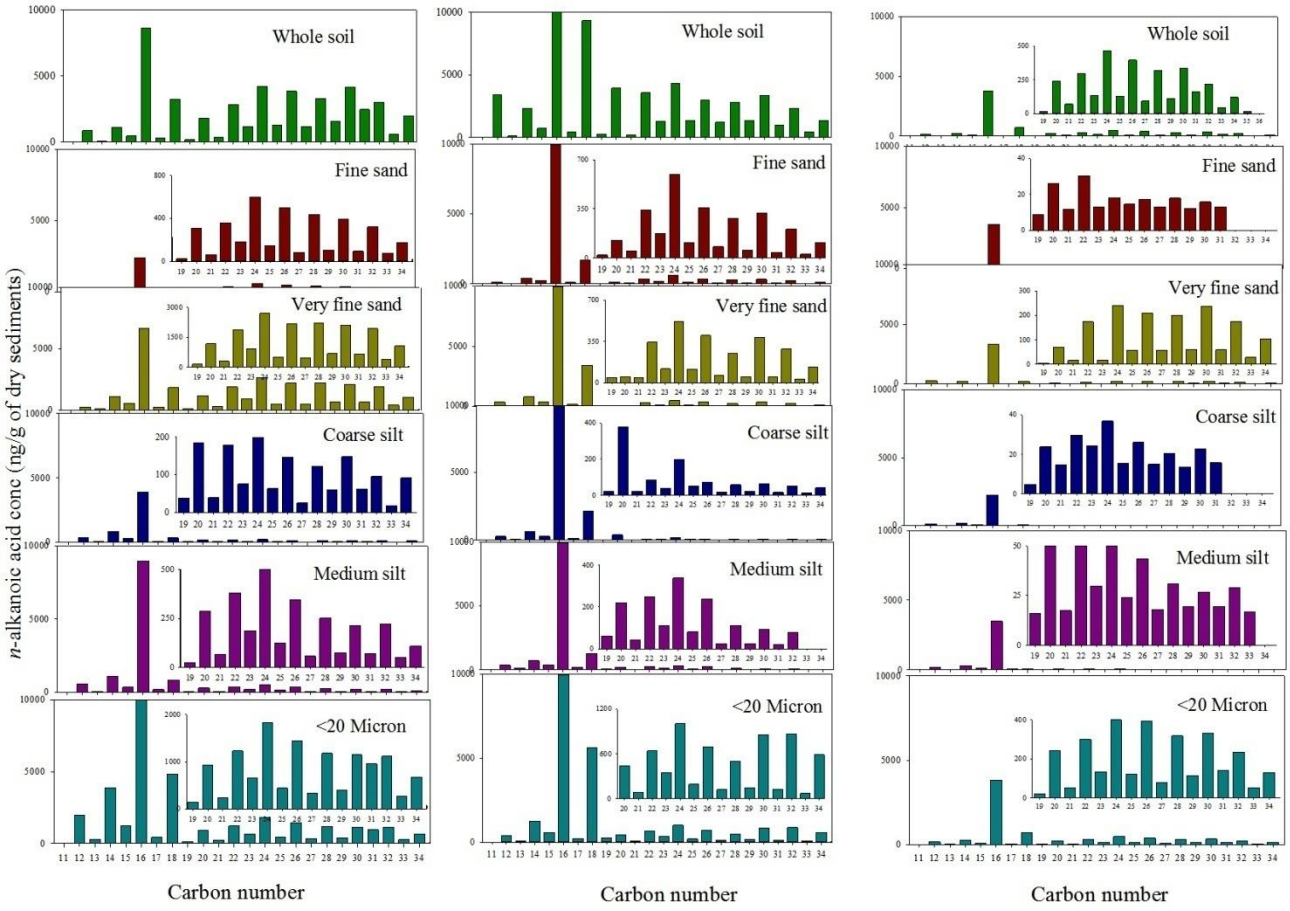


Figure 7: Abundances of the C_{12}–C_{34} *n*-alkanoic acids (in ng *n*-alkanoic acids/g dry soil) and the resulting *n*-alkanoic acids distribution patterns in the forest, grassland and mixed system soils (left, middle and right, respectively). The inset images illustrate the results for the *n*-alkanoic acids chain lengths (n-$C_{19–34}$) at a scale adapted to their low concentrations.

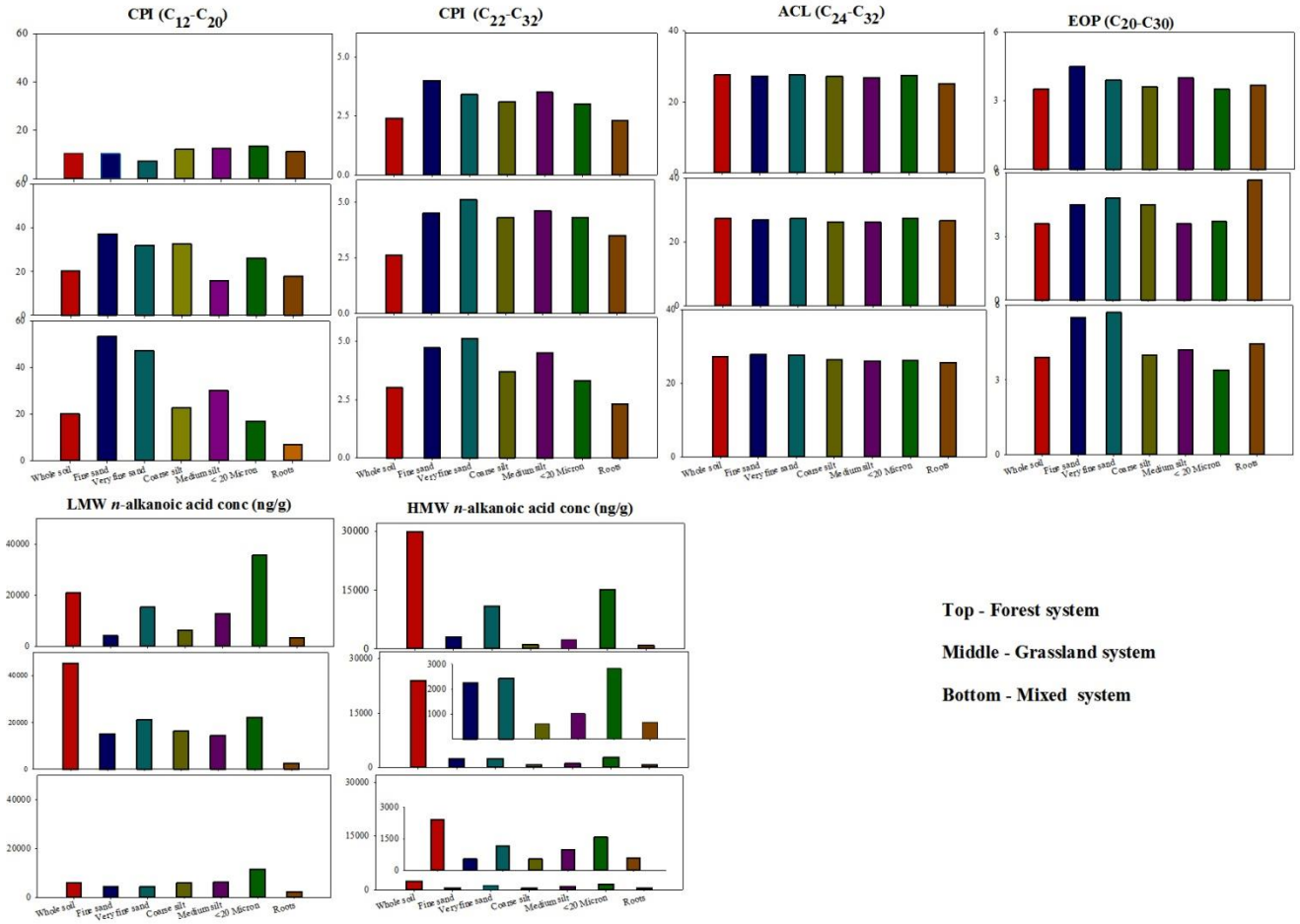


Figure 8: *n*-Alkanoic molecular indices and the relative proportions of the short and long-chain *n*-alkanoic acids for the investigated soil systems (Top – Forest system, Middle – Grassland and Bottom – Mixed system). The inset images showing long-chain *n*-alkanoic acids from grassland and mixed systems at a scale adapted to their low concentrations.

Table 2 a: Concentrations and other parameters of *n*-alkanoic acids in forest soil

Soil	Total al-acid	$CPI_{(22-32)}$	$CPI_{(12-20)}$	$ACL_{(24-32)}$	$EOP_{(20-32)}$	LMW	HMW	LMW/HMW
Whole soil	50964.3	2.4	10.4	27.7	3.5	21031.1	29933.1	0.7
Fine sand	7114.9	4.0	10.3	27.3	4.5	4175.7	2939.1	1.4
Very fine sand	30475.3	3.4	7.30	27.6	3.9	15351.2	15124.1	1.0
Coarse silt	7384.6	3.1	12.1	27.2	3.6	6338.3	1046.3	6.0
Medium silt	15165.5	3.5	18.4	26.9	4.0	12926.6	2238.8	5.7
<20 Micron	46492.0	3.0	13.3	27.4	3.5	35715.8	10776.2	3.3
Roots	4168.7	2.3	11.1	25.1	3.7	3304.7	864.0	3.8

Table 2 b: Concentrations and other parameters of *n*-alkanoic acids in grassland soil

Soil	Total al-acid	$CPI_{(22-32)}$	$CPI_{(12-20)}$	$ACL_{(24-32)}$	$EOP_{(20-32)}$	LMW	HMW	LMW/HMW
Whole soil	69068.6	2.6	19.9	27.5	3.6	45165.1	23903.4	1.8
Fine sand	17259.4	4.5	37.1	27	4.5	15008.1	2251.2	6.6
Very fine sand	23550.0	5.1	31.9	27.4	4.8	21113.8	2436.1	8.6
Coarse silt	16884.6	4.3	36.5	26.3	4.5	16278.5	1606.1	6.8
Medium silt	15285.6	4.6	15.7	26.2	3.6	14272.9	1012.7	14.0
<20 Micron	25068.8	4.3	26.1	27.4	4.4	22252.9	2815.9	7.9
Roots	3309.0	3.4	17.9	26.7	5.5	2641.4	667.6	3.9

Table 2 c: Concentrations and other parameters of *n*-alkanoic acids in mixed systems soil

Soil	Total al-acid	$CPI_{(22-32)}$	$CPI_{(12-20)}$	$ACL_{(24-34)}$	$EOP_{(20-32)}$	LMW	HMW	LMW/HMW
Whole soil	8270.5	3.0	20.4	27.2	3.9	5865.4	2405.1	2.4
Fine sand	6017.3	4.7	53.5	27.8	5.5	4456.2	1561	2.8
Very fine sand	5871.6	5.1	47.3	27.7	5.0	4444.5	1427	3.1
Coarse silt	6210.5	3.7	22.7	26.4	4.0	5689.9	520.6	10.9
Medium silt	6599.2	4.5	30.0	26.1	4.2	6077.5	521.7	11.6
<20 Micron	12396.6	3.3	16.8	26.2	3.4	11439.9	956.7	11.9
Roots	2721.3	2.3	6.71	25.6	4.6	2164.4	556.8	3.8

CHAPTER 5

Discussions

5.1 Grain-size and mineralogical distribution

Under similar climate conditions, soil forming intrinsic factors such as grain size distribution and mineral content, could be responsible for the transformation and accumulation of OM pools (Burke et al., 1989; Kögel-Knabner et al., 2008). From our dataset, forest, grassland and mixed systems have sand, silt and clay of similar content (Table 2). This indicates that grain size distribution did not have any specific role in the *n*-alkyl lipid abundance and distribution patterns in whole soil as well as in size fractions across the forest, grassland and mixed systems.

The dominant minerals assemblages found in these three systems are kaolinite, illite, quartz and feldspars. However, the variation in the abundance and distribution patterns of *n*-alkyl compounds across size fractions between forest, grassland and mixed systems put forward that the mineralogical compositions did not influence the OM distribution in these soils. Although climate and soil texture are the primary regional controls of the accumulation of OM, their influence on the distribution of lipids across size fractions, may be eclipsed by the effects of vegetation cover. In this context, it is suggested that vegetation cover determines the soil environment and play a key role in the functionality of the microbial communities present in these systems and can have significant impact on the lipid distribution pattern in the soil where the climatic conditions, chemical and physical characteristics of soil are similar.

5.2 Particle size fractionation

Physical separation of soil into five different organo-mineral fractions are used to identify and quantify the different stabilized lipid pools associated with sand, silt and finer fractions. Lipids bound to mineral grains of different size would help in understanding the pathways of OM associated with the mineral surface, either through directly sorbing the unmodified plant components or sorb to mineral surfaces after undergoing microbial transformation (Sokol et al., 2019).

The largest lipid content in the finer fraction of ~ 90% with respect to whole soils indicates that the finer fractions contribute the highest amount of lipid content in soils. The finer fraction with highest

lipid content differs from the sand and silts in terms of mineralogy and typically contains clay minerals (Hillel, 2012). The decrease in the total lipid contents related to the whole soil from finer fractions to sand and silt indicates a decrease in storage of lipids. Hence, we infer that the allocation of lipids in soil strongly depend on the adsorption of the compounds to clay minerals.

The possible explanation for the substantially higher in total *n*-alkyl content of the particle size fractions than the *n*-alkyl lipid content of the respective whole soils could be that the *n*-alkyl compounds may be trapped within the OM matrix or microaggregates. Disruptions of intact soil macroaggregates and the OM-mineral associations during fractionation/sieving process may have released these compounds from the soil matrix that otherwise may have remained locked inside aggregates in the whole soils.

5.3 *n*-Alkane abundance and distribution patterns of whole soils and particle size fractions

The distribution of *n*-alkane in the whole soils of all the systems reflected those of leaf lipids from higher terrestrial plants (Eglinton and Hamilton, 1967), suggesting that the long-chain *n*-alkanes compounds have not been affected considerably by diagenesis. Higher total *n*-alkanes content in the forests are in good agreement with findings of higher *n*-alkanes abundance in soil under forest than grassland and (Li et al., 2018). This indicates the occurrence of more plant litters accumulated in the forest that can effectively contribute to topsoil lipid content (Cristiane et al., 2010). Therefore, it is feasible that the accumulation of litter layers over time could lead to higher *n*-alkanes concentrations in the sub-soils. Moreover, thick litter assemblages in the forest soil could lead to anoxic conditions for higher preservation of plant derived OM materials and inhibit the decomposition of lipid compounds by microorganisms (Khatoon et al., 2017). This is further reflected by the occurrence of higher $HMW_{alk(23-33)}$ compounds with significant predominance of odd over even carbon number in the forest soil than grassland and mixed systems. Low reported total *n*-alkanes concentration in grassland compared to the forest, also reflected in high abundance of $LMW_{alk\ (11-22)}$ compounds, which we ascribe to the ongoing modification of plant-derived *n*-alkanes and contribution from microbial communities where the fine roots play a major role in proliferation of the microbial growth that influences the degradation of lipids in the soil (Cheng et al., 1990; Mason et al., 2005; Sokol et al., 2019). Our mixed system soils have very low *n*-alkanes concentrations, which we interpret to the little amount of litter samples at these sites and as degradation occurs, subsequently this could lead to very low concentration of *n*-alkanes in the soil.

These differences in *n*-alkanes concentration among forest, grassland and mixed systems could result from distinct plant wax abundance, chain length distribution, growth habits and physiology that may have influenced the *n*-alkanes signals in soils. Moreover, concentrations of *n*-alkanes in their vegetation

sources, land use system and effect of soil microorganisms must be accounted for, otherwise selective preservation or degradation cannot be identified or quantified when comparing between different terrestrial systems (Mueller et al., 2013; Freimuth et al. 2017).

The domination of C_{29} in forest, and the relatively higher concentrations of C_{33} in the grassland and mixed system soils, allow for differentiation between the OM input of trees and grass vegetation. However, care has to be taken when interpreting past vegetation based on the dominance of one n-alkanes compound over the others (Cmax), because this can lead to an underestimation of the root input, where the leaves were not the main contributors to the topsoil (Jackson et al., 1996). Our data from roots *n*-alkanes distribution patterns shed some light about the dominance of *n*-C_{31} than *n*-C_{29} in forest (Figure 6a). Therefore, the roots can potentially affect the distribution of *n*-alkanes by imprinting a new signal over that of the leaf. However, roots contain a much lower concentration of *n*-alkanes and are not likely to be the primary source of *n*-alkanes in these systems (Table 1a; 1b; 1c).

The *n*-alkane patterns in size fractions of forest, grassland and mixed systems vary in the relative abundance of higher vs. lower compounds. Relatively higher concentration of $HMW_{alk(23-33)}$ compounds with pronounced odd-over-even *n*-alkanes carbon number in sand- than silt-size and finer fractions implies that the OM accumulated in sand is dominated by relatively fresh or partly decomposed plant inputs (Six et al. 2004). Quartz, which is the dominant mineral in sand fractions, has less capability to hold the OM due to small surface area and thin clay coating compared to clay minerals (Figure 5a). This signifies the occurrence of temporal pools of *n*-alkane in the sand fractions (Lichtfouse et al., 1998; Cayet et al., 2001). Moreover, higher concentration of $HMW_{alk(23-33)}$ compounds in sand-size apart from fresh input plant material, suggests low modification of plant-derived compounds where low nutrient availability in these fractions account for low microbial biomass (Sessitsch et al., 2001). The lowest total *n*-alkanes concentration and lowest $HMW_{alk(23-33)}$ compounds in silt-size fractions from forest, grassland and mixed systems indicates the poor preservation and modification of the long-chain *n*-alkanes in these fractions of soils. This denotes the higher microbial biomass found in silt-size fractions (Sessitsch et al., 2001; Selesi et al., 2007; Hemkemeyer et al., 2018). This finding further demonstrates that microbial communities associated with particle size can greatly affect the distribution of lipid compounds in the soil. Highest total *n*-alkane concentration, $HMW_{alk(23-33)}$ and $LMW_{alk(11-22)}$ compounds in finer fractions than sand and silt in forest, grassland and mixed system denotes the interaction of *n*-alkanes with the minerals present in finer fractions, mainly by chemical adsorption to clay mineral surfaces (Quenea et al., 2004). In addition, the occurrence of $LMW_{alk(11-22)}$ and $HMW_{alk(23-33)}$ compounds with weak odd-over-even predominance in substantial amounts reflects the contribution of modified plant biomass and contribution from the microbial communities in the finer fractions.

The short-chain *n*-alkanes compounds (C_{11}-C_{21}) with no odd or even carbon number predominance in sand-, silt-size and finer fractions of all the systems could be attributed to the presence of modified plant materials or a microbial input in the soil (Wiesenberg et al., 2004). Short-chain *n*-alkanes are known to be more sensitive to degradation by microorganisms than the long-chain *n*-alkanes compounds (Moucawi et al., 1981; Amblès et al., 1993). As a result, higher abundance and distribution patterns of short-chain *n*-alkanes in the finer fractions likely indicates the potential of clay mineral surfaces bound with the finer particle fractions in influencing the preservation of *n*-alkanes compounds derived from microbial origin with respect to sand and silt (Moucawi et al., 1981; Quenea et al., 2004) (Figure 5b).

5.3.1 Molecular proxies of *n*-alkanes

In addition to the difference in the *n*-alkane concentration across the investigated soil and size fractions under different vegetation cover; our results indicate that the distribution patterns vary in the abundance of odd vs. even carbon number long-chain *n*-alkanes. A comparatively higher $CPI_{alk(23-33)}$ value in forest than grassland and mixed systems confirmed the effective preservation of plant-derived *n*-alkanes and has not significantly altered by diagenesis compared to grassland and mixed systems. This indicates that the relative abundance of *n*-alkanes chain length varied with soil under different vegetation cover. More conservation systems like forest presented an increase in the preservation of long-chain *n*-alkanes that can have a positive influence on the dynamics and stock of soil carbon (Cristiane et al., 2010).

A decreasing trend in long-chain *n*-alkanes $CPI_{alk(23-33)}$ values with decreasing grain size from the sand to the finer fraction in these systems indicate the occurrence of plant-derived compounds in the sand fractions (Figure 5a). A possible modification of plant-derived *n*-alkanes by microbial activity in silt and finer fractions may have contributed to the even-carbon number long-chain *n*-alkanes. Organo-mineral associations in these fractions show a microbial fingerprint that reflected the low $CPI_{alk\ (23-33)}$ values, suggests the high affinity of microbes to small size fractions of soil (Grimalt et al., 1998; Sessitsch et al., 2001; Selesi et al., 2007; Brittingham et al., 2017).

The percentage of decrease (~60%) in $CPI_{alk(23-33)}$ from sand to finer fractions is much more pronounced in the forest system than the decreasing pattern of $CPI_{alk(23-33)}$ from sand to finer fractions in grassland (~50%) and mixed systems (~30%). Thus, we infer that relative abundance of odd/even carbon number long-chain *n*-alkanes is influenced by the types of aboveground vegetation cover which can substantially alter the compositions, diversity and functions of soil microbial communities. Soil microorganisms are an important bioactive component in the terrestrial ecosystems. Therefore, these

results observed in the decreasing pattern of $CPI_{alk(23-33)}$ that vary between forest, grassland and mixed system suggest that the difference in litter composition of the type of vegetation could lead to changes in the nutrient content in the soil. Moreover, vegetation cover defines the root systems that can alter the soil conditions (Prescott et al., 2013; Moscatelli et al., 2018). The soil conditions affect the ability and function of microorganisms to utilize the *n*-alkanes compounds. Thus, the microbial functionality and diversity differ in different terrestrial ecosystems (Anderson et al., 2009; Weng et al., 2021).

Fungal network in regions of higher microbial density such as grassland topsoil, promotes the translocations of the plant-litter away from the litter surface leading to the formation of persistent mineral associated OM aggregates (Sokol et al., 2019). It is known that the microbes associated with the finer fractions, released polysaccharides that serve as the binding material between the mineral grains and the plant-derived OM (Sollins et al., 1996). This emphasized the important role of microorganisms in incorporation of plant-derived OM-mineral aggregates in the finer fractions.

Short-chain *n*-alkanes having roughly constant $CPI_{alk(11-21)}$ values (<1) in whole soils from all the systems without significant odd/even carbon number predominance implies microbial origin (Cranwell, 1981; Meyers and Ishiwatari, 1993). In different size fractions of forest, grassland and mixed system, low $CPI_{alk(11-21)}$ in silt-size and finer fractions than sand fractions represent diagenetic modification of the short-chain compounds corroborating the observation from long-chain $CPI_{alk(23-33)}$.

Higher $ACL_{alk(25-33)}$ in grassland suggest the OM sources with leaf lipid that derived from grasses have higher values than that of the trees (Cranwell, 1973). This reflects that the vegetation cover is the main influence on the length of the carbon chain *n*-alkanes that belongs to the terrestrial plant lipids. Minimal difference in $ACL_{alk(25-33)}$ across size fractions with slightly higher value in sand and decrease with decreasing size in forest (~1.6%), grassland (~2%) and mixed system (~1%) suggests a role of post-depositional processes in controlling the dominant chain length distribution of *n*-alkanes compounds. Notably, the $CPI_{alk(23-33)}$ is weakly correlated with the $ACL_{alk(25-33)}$ indicating that the ratio of odd carbon number long-chain *n*-alkanes is not affected by the changes in the odd vs. even carbon number long-chain *n*-alkanes across the grain size fractions. This suggests the role of post depositional process.

Lower C_{29}/C_{31} in grassland whole soil seem plausible that there is greater predominance of longer chain lengths *n*-C_{31} alkanes in grasses than in trees. C_{29}/C_{31} cannot serve as generalized proxies for separating *n*-alkanes derived from grasses and mixed systems.

Increases in C_{29}/C_{31} with decreasing size fractions demonstrate that the ratios of these two compounds provide insights on the relative abundance of *n*-alkane chain length in different size fractions of soil. Higher C_{29}/C_{31} values in silt size in forest and grassland compared to sand and finer fractions signifies more pronounced biodegradation of odd carbon number long-chain *n*-alkanes compounds with

decreasing size fraction. This substantial degradation of long-chain compounds corroborates with the fact that larger microbial biomass inhabit in smaller size fractions (Sessitsch et al., 2001).

Grassland and mixed systems have lower $OEP_{alk(27-33)}$ than the forest, suggesting higher ongoing modification of long-chain *n*-alkanes in soil under grasses and mixed systems compared to soil in forest. These observations attributed to the specificity of grassland and mixed systems. Higher fine root density in grassland and mixed system soils than that of forest allows more infiltration of water and air leading to oxic conditions and less preservation of OM contributed by the dominant vegetation in the grassland and mixed systems (Chikaraishi and Naraoka, 2006; Khatoon et al., 2017). Higher microbial biomass is usually found from the rhizosphere soil, which is the narrow region of soil that is directly influenced by roots, than the other parts of the soil horizon where roots are less abundant (Cheng et al., 1990; Sokol et al., 2019). Hence, this highlights that microbial activity is relatively intense in grassland and mixed systems compared to the forest system. High $OEP_{alk(27-33)}$ in forest results from the thick layer of litter accumulated in soil under trees that can block the infiltration of air and water which lead to anoxic conditions. Microbial activity is known to be favorable under oxic conditions and abruptly decreased under anoxic conditions (Dyckmans et al., 2006). Hence, low microbial activity in soil under forest favored the preservation of plant derived *n*-alkanes that effectively contribute to topsoil long chain *n*-alkanes content that reflect in high OEP.

Decrease of $OEP_{alk(27-33)}$ from sand-size to finer fractions of forest, grassland and mixed systems is correlated with the $CPI_{alk(23-33)}$ and C_{29}/C_{31}, an indication in the decrease in abundance of odd carbon number long-chain *n*-alkanes compounds with decreasing grain size.

Lower $OEP_{alk(27-33)}$ in the finer fractions of all the three systems reflects the influence of biodegradation of odd carbon number long-chain *n*-alkanes by the microorganism associated with these fractions.

Decrease of $OEP_{alk(27-33)}$ from sand-size to finer fractions is more pronounced in grassland (~60%), compared to the forest (~54%) and mixed system (~52%). Grasslands, which have higher fine root biomass compared to forests and mixed systems could have higher microbial growth associated with the mineral grains in finer fraction of soil. Therefore the degradation patterns of *n*-alkanes were different for the three systems and more intense in size fractions from grassland. This observation indicates that the trends are influenced by differences in environmental properties such as the diversity in microbial communities and the soil microclimate (Chikaraishi and Naraoka, 2006; Dyckmans et al., 2006; Khatoon et al., 2017).

5.4 *n*-Alkanoic acids abundance and distribution patterns of whole soils and particle size fractions

The occurrence of large abundance of $LMW_{acid(12-20)}$ compounds in grassland is likely caused by the modification of long-chain *n*-alkanoic acids to short-chain compounds in the soil and by the input of microbial communities (Otto et al., 2005; Sokol et al., 2019). Higher abundance of $HMW_{acid(22-34)}$ compounds with even carbon numbers in forest compared to grassland and mixed systems indicate low to moderate levels modification of long-chain *n*-alkanoic acid in the forest soil.

In size fractions, abundance of $HMW_{acid(22-34)}$ compounds in sand is not expectable, as the plant-derived *n*-alkanoic acids is more likely to degrade at a faster rate than the *n*-alkanes (Wiesenberg et al., 2004). We infer the persistence of plant-derived *n*-alkanoic acids, contained in sand-size fractions of soils from these three systems, may have accumulated in sand-sized aggregates (>63μm) (Angst et al. 2018). Accessibility of *n*-alkanoic acids to microbes may have decreased by occlusion within the microaggregates. Small size pores could restrict the entry of microorganisms and low oxygen levels limited to microbial activity (Sollins et al., 1996). Intense microbial activity in silt fraction is reflected by much lower $HMW_{acid\ (22-34)}$ observed in silt fractions than sand and finer fractions of all the systems. These observations highlight the variation in microbial communities associated with differently sized soil particles (Grimalt et al., 1998; Sessitsch et al., 2001; Selesi et al., 2007; Brittingham et al., 2017).

Lipids with carboxyl groups were known to be preferentially enriched in clay minerals such as phyllosilicates (Jones et al., 2014; Ghosh et al., 2019). We infer that the nature of functional groups of *n*-alkanoic acids could result in the higher abundance of total *n*-alkanoic concentration, $HMW_{acid(22-34)}$ and $LMW_{acid(12-20)}$ compounds in the finer fractions of all the three systems. Regardless of the differences in vegetation cover, another reason for the preservation of these compounds in the finer fractions could result from occlusion within the microaggregates reduces accessibility to the microbes (Lützow et al., 2006; Clemente et al., 2011; Yang et al., 2020). Although *n*-alkanoic acid is considered to be relatively labile compared to *n*-alkanes, affinity to clay minerals could be an important pathway in the preservation of these compounds particularly in silt-size and finer fractions across the different ecosystems (Figure 5b).

5.4.1 Molecular proxies of *n*-alkanoic acids

$CPI_{acid(22-32)}$ values in forest, grassland and mixed system soils point to a dominant plant wax source. Unlike *n*-alkanes, small differences in $CPI_{acid(22-32)}$ values in these systems cannot differentiate the specific vegetation source. Our results show that the amount of *n*-alkanoic acids between the systems vary in the concentration of higher molecular weight compounds (sum of even and odd carbon) which did not reflect in the ratio of long-chain even-over-odd *n*-alkanoic acids ($CPI_{acid(22-32)}$).

Comparatively higher $CPI_{acid(22-32)}$ in sand-size fractions than in silt and finer fractions allow us to identify the source of *n*-alkanoic acids with characteristic of plant-derived and contribution of microbial input in sand and finer fractions respectively. Small changes in $CPI_{acid(22-32)}$ among sand, silt and finer fractions in all the systems, suggests the processes/mechanisms that involved for the preservation of the long-chain even carbon number *n*-alkanoic acids leading to small changes in $CPI_{acid(22-32)}$ across the soil size fractions. These long-chain even compounds can be selectively preserved in the silt and finer fractions with little or no alteration by adhering with the clay mineral surfaces. The adheration of *n*-alkanoic acids to the clay mineral surfaces is determined by their functional groups (COOH). (Quénéa et al., 2004 ; Lützow et al., 2006; Kleber et al., 2015; Angst et al., 2018; Yang et al., 2020). These compounds in finer fractions that have been protected from the microbial activity can remain in the soil up to several years to decades and centuries.

Comparison between the different systems, lower % of decrease in $CPI_{acid(22-32)}$ from sand to finer fractions in grassland (~15%) when compared to forest (~25%) and mixed system (~35%), indicate the persistence of long-chains even carbon number *n*-alkanoic acids in finer fractions as well. This suggests another possibility of *n*-alkanoic acid stabilization mechanisms where different microbial communities present in different ecosystems or the microbial activities associated with the particle size have different affinity towards *n*-alkanoic acids for their metabolic activities. Thus, soil microorganisms mediate turnover of *n*-alkanoic acids to cover their nutrient and energy needs for maintenance and growth (Sessitsch et al., 2001; Dyckman et al., 2006; Schmidt et al., 2011; Miltner et al., 2012; Hemkemeyer et al., 2018).

Therefore, the concept of microbial roles in *n*-alkanoic acid stabilization by understanding the dynamics across the grain size fractions is key aspects to the function of soils in the global carbon budget. However, the precise nature of the interactions between these lipids compounds, soil microorganisms needs further study as this can have important implications on OM turnover. Our results $CPI_{acid(22-32)}$

demonstrate that the effects of vegetation cover and microbial interactions on the distribution of long-chain n-alkanoic acids are more noticeable in size fraction than in whole soil.

$CPI_{acid(12-20)}$ derived from short-chain *n*-alkanoic acids is higher than that of long-chain $CPI_{acid(22-32)}$. Higher values in short-chain *n*-alkanoic acid $CPI_{acid(12-20)}$ with respect to long-chain $CPI_{acid(22-32)}$ observed in whole soils and different size fractions, implies that the even carbon number *n*-alkanoic acids are present in higher quantities in the short-chain (C_{12}-C_{20}). One possible explanation for this could be that the short-chain *n*-alkanoic acids already synthesized in plant leaves and with the contribution from microbial communities result in higher abundance of these compounds in the soil. Studies have shown that short-chains *n*-alkanoic acids are generally distributed in small-sized aggregates. Small-size aggregates provide long-term protection that allows for short-chains to be more stable than long-chain acids (Six et al., 2004, 2000; Lutzow et al., 2006; Yang et al., 2020). Much higher value of $CPI_{acid(12-20)}$ in all the five size fractions than that in their respective whole soils, mostly in the grassland and mixed system indicates that the short-chain *n*-alkanoic acids has been trapped within the microaggregates and bind with the soil OM matrix. Disruptions of soil during the soil size separation process may have released these compounds that may otherwise remain locked inside aggregates in the whole soils.

Higher $CPI_{acid(12-20)}$ in sand-size than in finer fractions for grasslands and for mixed systems soils are consistent with the suggestion of admixture from terrestrial plant and microbial source, whereas much higher contribution from microbial communities in the finer fractions result in lower $CPI_{acid(12-20)}$. The similar behavior of $CPI_{acid(12-20)}$ for forest soil across grain size accounts for less influence from microbial activity in the forest soil that does not drastically alter the even-over odd predominance in the short-chain *n*-alkanoic acids in the finer fractions.

From the present study, it was noticed that there was no considerable changes in the $ACL_{acid(24-32)}$ among the whole soils and across the size fractions of forest, grassland and mixed systems which signifies a similar abundance of long-chain odd carbon number *n*-alkanoic acids across the size fractions. However the *n*-alkanoic acid concentration is different between forest, grassland and mixed systems as well as across size fractions which did not reflect in the chain length ratio ($ACL_{acid(24-32)}$). This limits the use of $ACL_{acid(24-32)}$ in past vegetation reconstructions and in studying the distribution patterns of *n*-alkanoic acids to identify changes in vegetation cover.

Additionally, no considerable variation in $EOP_{(22-32)}$ between the three systems suggests that regardless of the type of overlying vegetation cover with different total *n*-alkanoic acid concentrations in soil, *n*-

alkanoic acid did not reflect the changes in even vs. odd predominance between the systems. This indicates that long-chain *n*-alkanoic acid stages of modification are indistinguishable between the forest, grassland and mixed system.

However, a slight decreasing trend in $EOP_{(22-32)}$ from sand size to finer fractions of forest (~22%), grassland (~25%) and mixed systems (~32%) signifies that the even carbon number long-chain *n*-alkanoic acids are present in slightly higher amounts in coarser fractions. The changes in $EOP_{(22-32)}$ between the size fractions of these soils therefore point to a low to moderate level of modification due relatively large contribution of *n*-alkanoic acids derived from higher plant waxes in the sand fraction and a result of microbial alteration in finer fractions.

The odd carbon number *n*-alkanoic acids observed in the finer fractions particularly grassland and mixed systems, probably originate from the reworking by microbial activity of the even carbon number long-chain *n*-alkanoic acids through the β-oxidation and subsequent chain shortening (Dinel et al., 1990; Quenea et al., 2005; Yang et al., 2020). Greater microbial activity in the grassland and mixed systems, as compared to forest systems suggests presence of different microbial communities in different systems.

5.5 Mechanisms of stabilization of *n*-alkyl lipids

Understanding the association between *n*-alkyl lipids and different particle size fractions is important for assessing possible mechanisms that control the stabilization of these compounds in the soil. This may contribute to future model development and ultimately for decision making relative to carbon sequestration. The amount and distribution patterns of *n*-alkyl lipids in any soil system could represent the net result of the quality and quantity from the plant litter added to soil, microbial interaction and extent of degradation.

Stabilization of *n*-alkanoic acids:

Soil microorganisms possess the enzymatic capability to degrade the *n*-alkanoic acids. The decomposition of these compounds in soils is controlled by β-oxidation, a process in which *n*-alkanoic acids are decomposed to CO_2 and H_2O (Dinel et al., 1990; Rustan and Drevon, 2005). However the CPI values of long chain *n*-alkanoic acids did not reveal such a process, otherwise we would have observed a low CPI value with decreasing grain size. CPI did not reflect the significant changes in even-over-odd ratio from sand to finer fractions, whereas the abundance of HMW *n*-alkanoic acid varies across size fractions.

Our results suggest that the processes leading to small differences in CPI could be related to the nature of *n*-alkanoic acid functional groups. The carboxyl groups of *n*-alkanoic acids can interact with clay

mineral through ligand exchange and polyvalent cations bridges while the alkyl chains promote hydrophobic zones (Kleber et al., 2015; Lutzow et al., 2006). In addition, *n*-alkanoic acids can be stabilized by occlusion within microaggregates (Angst et al., 2018; Lehmann et al., 2007). Therefore, the interaction of these compounds with clay mineral surfaces could lead to the stabilization of these compounds that are more labile that can be easily degraded by microorganisms and ultimately, for the regulation of soil OM preservation.

Another possible reason for the stability of long-chain even carbon number *n*-alkanoic acids in silt and finer fractions could result from the affinity/requirements of microorganisms towards *n*-alkanoic acids for their metabolic activities. Moreover, specific microbes associated with particle size are probably another possible reason for the distribution of *n*-alkanoic acid across the grain sizes (Moucawi et al 1981; Sessitsch et al., 2001; Hemkemeyer et al., 2018).

Stabilization of *n*-alkanes:

The slower turnover rate of *n*-alkanes than *n*-alkanoic acid was confirmed previously (Wiesenberg et al., 2008a; Griepentrog et al., 2015, 2016). The stability of *n*-alkanes in these investigated soils and across size fractions with respect to microbial degradation is determined by their molecular structure. The alkyl chain exhibits hydrophobic properties that cause water repellency. This could possibly promote the microbial attraction and influence decomposition by microorganisms.

n-Alkanes in soils can be degraded to *n*-methyl ketones through β-oxidation (Chaffee et al., 1986; Ambles et al., 1993), which allows us to trace *n*-alkanes degradation across different size fractions. Decreases in CPI values of long-chain *n*-alkanes from sand to finer fractions observed in forest, grassland and mixed system, reveal the process of oxidative reactions mediated by microorganisms with decreasing grain size.

5.6 Implication to paleoecological studies

Our results fall in line with previous findings from other studies that considered soil under different vegetation cover and showed information on lipids abundance in soil. Such other studies did not explicitly focus on the identification of incorporation and preservation pathways of *n*-alkanes and *n*-alkanoic acids across the grain size fractions of soil having different vegetation cover. Moreover, previous studies did not consider the changes incurred by the variation in vegetation cover under similar climatic conditions and similar intrinsic soil forming factors such as particle size distribution and mineral content. However, the data they do report is generally consistent with our findings in terms of relatively younger lipid compounds in sand-size fractions and older compounds of long term turnover time in the clay-size fractions (Quenea et al., 2004; Clemente et al., 2011).

Together with previous findings, our result suggests that the *n*-alkanes and *n*-alkanoic acids abundance and chain length patterns in whole soil and size fractions are associated with the dominant vegetation cover. Vegetation cover, in turn, defines the root system and litter surface, which affect the movement and bioavailability of air and water (Young et al., 2008).

n-Alkyl lipid molecular abundance and distribution patterns of whole soil and size fractions can be used as biomarkers in assessing changes past vegetation cover, as well as the occurrence of carbon stabilization mechanisms. The present study in distribution pattern of *n*-alkanes and *n*-alkanoic acids in soil size fractions from forest, grassland and mixed systems provide evidence for the compartmentalization of *n*-alkyl lipids that is of relatively fresh plant material in sand-size and contribution of microbial–derived compounds in finer-size fractions.

Additionally, results from *n*-alkanes molecular parameters such as CPI, ACL, OEP, ratio of lower molecular weight vs. higher molecular weight suggest that *n*-alkanes could be used to distinguish about the vegetation source in whole soil. These molecular indices derived from *n*-alkanes signal the influence of different types of vegetation cover in degradation patterns of molecular substance in different size fractions and identification of incorporation/preservation pathways of *n*-alkanes compounds in different size fractions. Their response to factors incurred by vegetation cover and microbiological interactions offer opportunities to apply these parameters derived from *n*-alkanes identify the input of OM in the soils that are quite useful in paleoecological studies.

Molecular indices from *n*-alkanoic acids (CPI, ACL and EOP) with no significant variability between the investigated soils of different vegetation cover illustrates how the knowledge from litter input and soil under different vegetation cover signals, does not directly reflect in the *n*-alkanoic acids record in whole soil and size fractions. Even though the faster turnover rates of *n*-alkanoic acids than the *n*-alkanes was confirmed previously (Wiesenberg et al., 2008), CPI derived from long chain *n*-alkanoic acids provide valuable information on long term carbon sequestration in the finer fractions of soil regardless of type of modern dominant vegetation cover. Nevertheless, the complex mechanisms for the preferential stability of long-chain even carbon number *n*-alkanoic acid in the silt and finer fractions was not clearly understood. However, the chemical structure of *n*-alkanoic acids (the nature of functional group) and different affinity/requirements of microbial community towards *n*-alkanoic acids for their metabolic activities that are presence in different terrestrial ecosystems as well as the specific microbes associated with particle size is probably the possible reason for the distribution of *n*-alkanoic acid across the grain sizes (Moucawi et al., 1981; Sessitsch et al., 2001; Hemkemeyer et al., 2018). Thus, more study is needed for the precise nature of the interactions between these compounds and microorganisms concerning the dynamics in soil OM turnover which play the important role in global carbon budget.

CHAPTER 6

Conclusions

We studied the *n*-alkanes and *n*-alkanoic acids in five different size fractions and compared their concentration and distribution patterns between the size fractions, with that of the whole soil and soil under different vegetation cover. Total *n*-alkyl lipid concentrations were inconsistent among different soil systems. HMW *n*-alkanes and *n*-alkanoic acids concentration were highest in the forest and LMW *n*-alkanes and *n*-alkanoic acids concentration were highest in grassland. Mixed systems however have the lowest HMW and LMW *n*-alkyl lipids concentration than the forest and grassland. The difference in the abundance of *n*-alkyl compounds among forest, grassland and mixed systems whole soils suggests that the vegetation cover plays a major role in lipid distribution in soil and little influence of climatic, soil textural and mineralogical aspects.

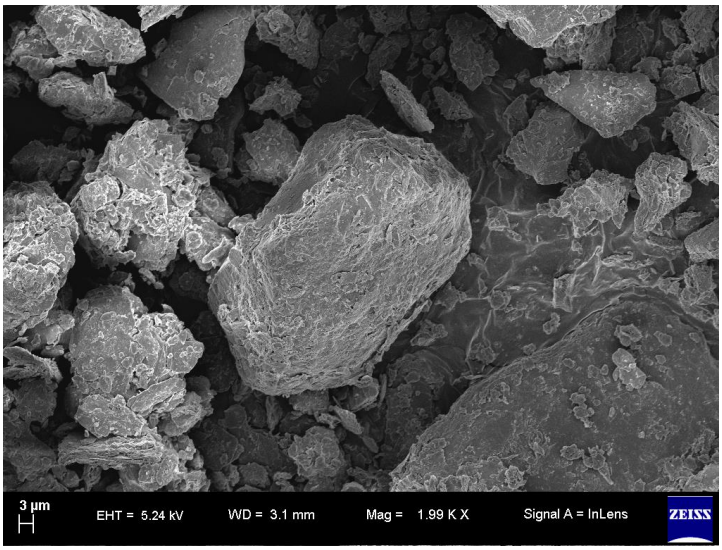
Particle-size fractionation provides valuable information that would strengthen the current understanding for evidence into the sources, degradation patterns and the preservation mechanisms of *n*-alkyl compounds in the soil. In five size fractions *n*-alkane and *n*-alkanoic acid concentrations were highest in the finer fractions. Abundance of HMW *n*-alkyl compounds with characteristics of plant-derived were more in sand fractions compared to silt and finer size fractions. Silt fractions have the lowest HMW *n*-alkyl compounds attributed to microbial diversity associated with these fractions. HMW and LMW *n*-alkyl lipids with lower odd/even predominance were highest in the finer fractions. The modification of plant derived and subsequent contribution from microbial communities resulted in an enrichment of short chain *n*-alkyl compounds and even long chain *n*-alkanes in finer fractions.

The present study in the distribution patterns of *n*-alkanes and *n*-alkanoic acids from forest, grassland and mixed systems, provide evidence on the incorporation and preservation pathways of lipids across size fractions. Several molecular parameters such as carbon preference index (CPI) and the average chain length (ACL) and odd-over-even predominance (OEP) could be used to differentiate the presence of plant OM in sand-size fractions and the contribution of microbial input in the finer fractions. More importantly, in size fractions, the fact that the decrease of CPI of long-chain *n*-alkanes from sand to finer fractions correlated significantly with the OEP (alkanes) suggests that the information on the incorporation pathways of plant-derived OM and microbial input in different size fractions stored in *n*-alkanes patterns. From this study, we could see the distribution patterns of *n*-alkanes in different size

fractions in different systems as an important paleoecological proxy to establish the interpretation in vegetation compositions.

The small changes in CPI, ACL and EOP derived from *n*-alkanoic acids between coarse and finer fractions did not reflect in the *n*-alkanoic acid concentrations. *n*-Alkanoic acids is known to have faster turnover rates with respect to other straight-chain lipids such as alkanes. The presence of the even carbon number long-chain *n*-alkanoic in the finer fractions, signal in CPI, is also useful to validate and strengthen the current understanding of conceptual organic carbon pools.

Interaction of these *n*-alkyl lipid compounds with the clay minerals, occlusion within microaggregates and biological process drive by the fungal hyphae are the major stabilization process. Therefore, evaluation of stabilized lipid pools in size fractions in terms of incorporation and preservation pathways of source OM, not addressed previously, would be a step forward for prediction of organic carbon response to future climatic changes.



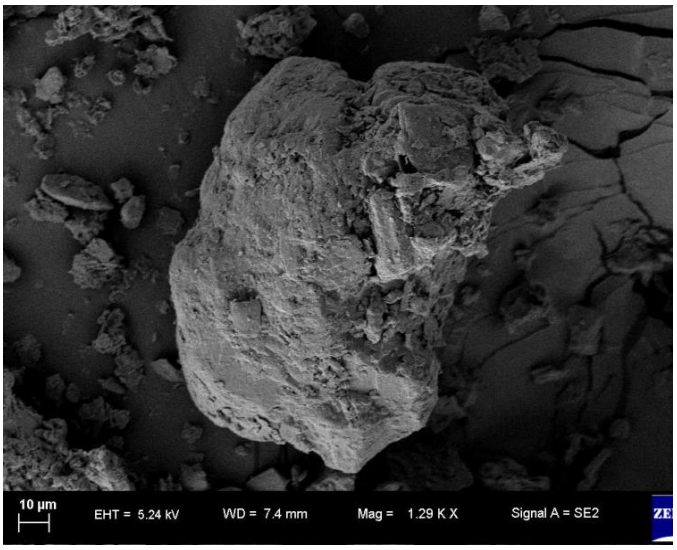
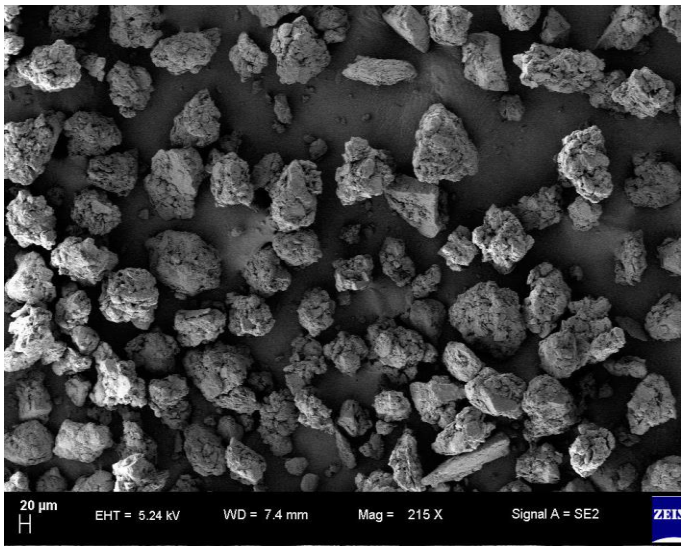


Figure 5a: Scanning electron microscopy (SEM) images of the sand-size fractions, showing the macroaggregates (3µm) (zoom out-left) and the sand grain (10µm) (zoom in-right).



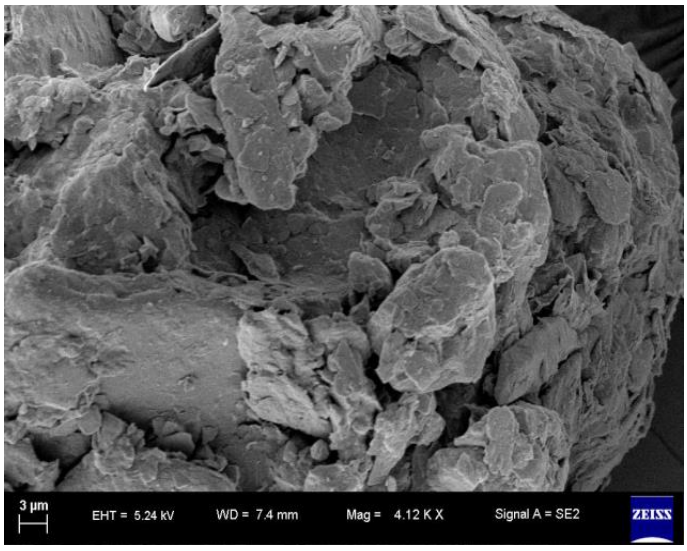


Figure 5b: Scanning electron microscopy (SEM) images of the finer fractions, showing the microaggregates (20µm) (Zoom out-left) and the clay mineral assemblages (3µm) (zoom in-right).

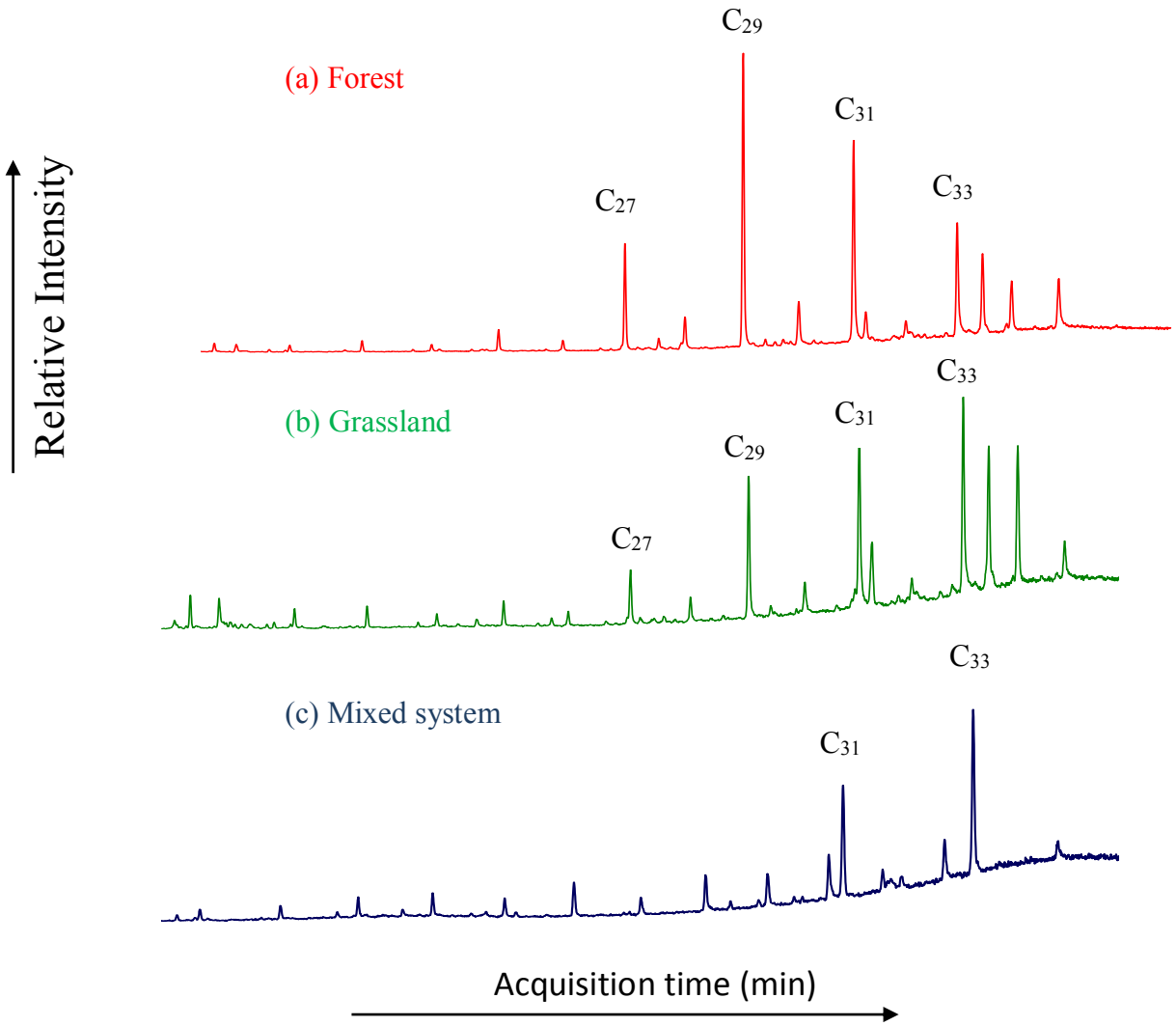


Figure 4a: GC-MS chromatograms of *n*-alkanes from (a) forest; (b) grassland and (c) mixed system, showing odd over even carbon predominance in higher molecular weight (HMW) *n*-alkanes. Numbers mentioned on the peaks refer *n*-alkanes carbon number.

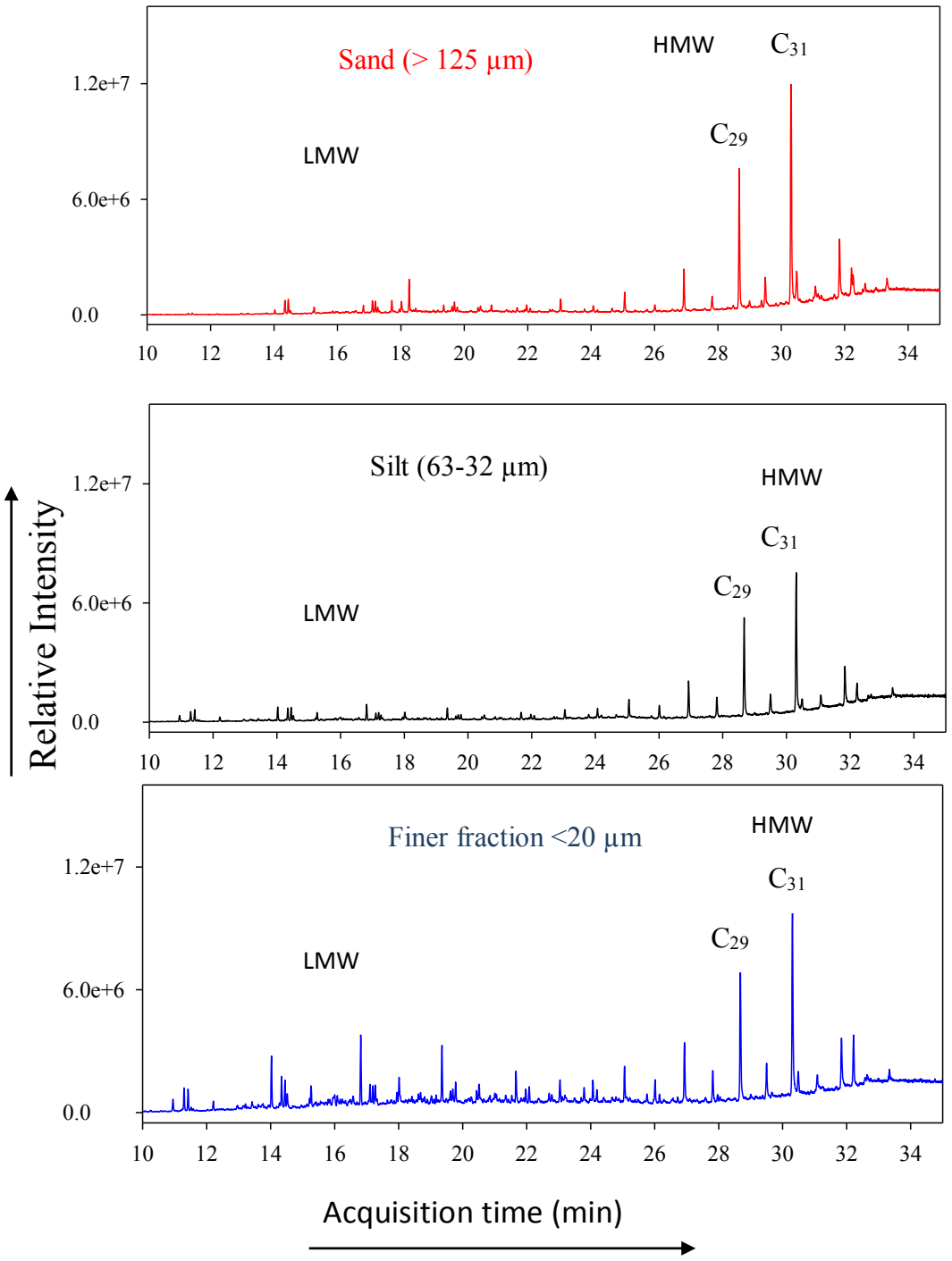


Figure 4b: Representative *n*-alkanes spectra and various chain-lengths of the (a) sand-, (b) silt- and (c) finer size fractions. Numbers mentioned on the peaks refer *n*-alkanes carbon number. LMW-Low molecular weight; HMW-High molecular weight compounds

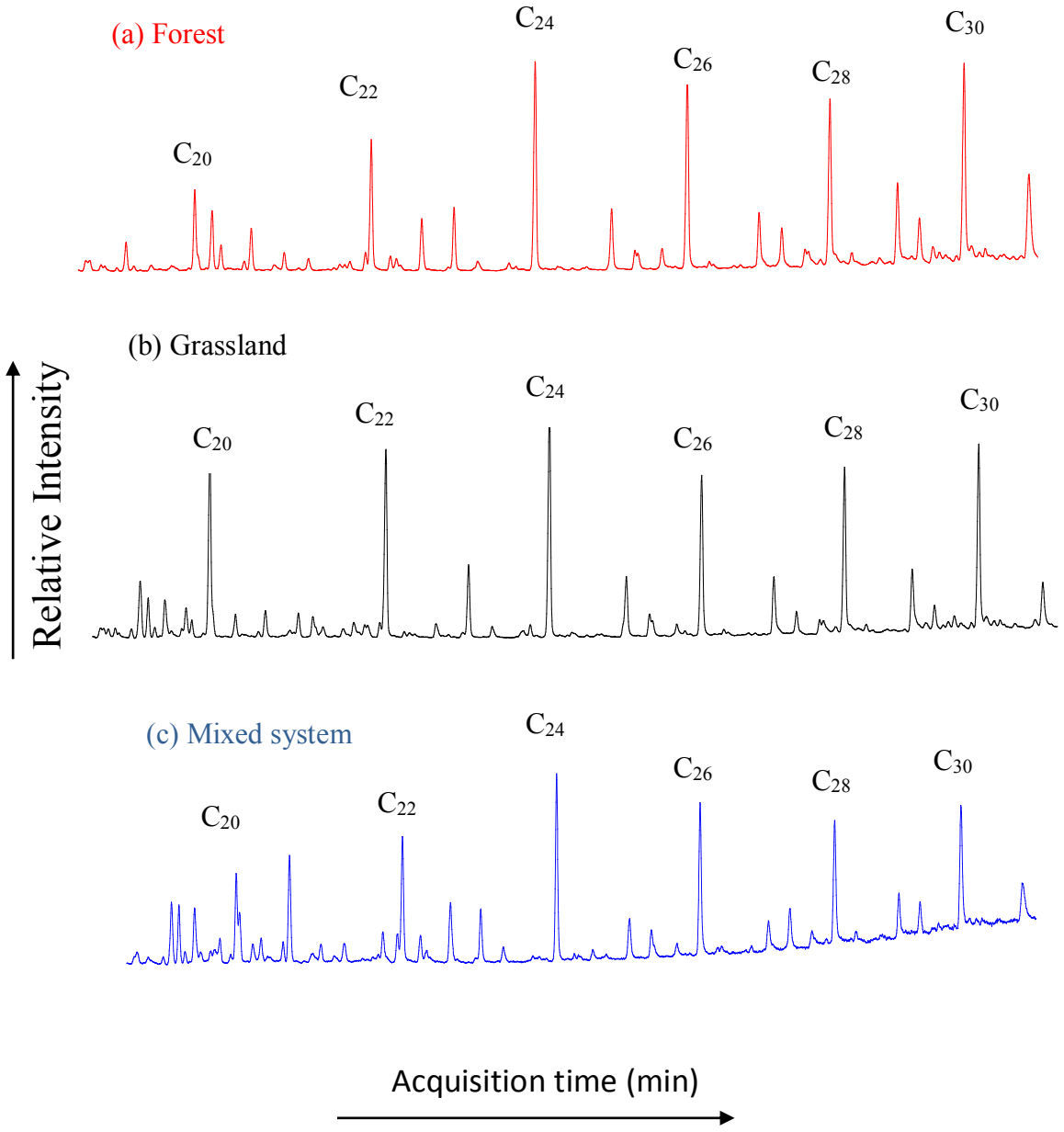


Figure 4c: GC-MS chromatograms of *n*-alkanoic acids from (a) forest; (b) grassland and (c) mixed system showing even over odd carbon predominance in the higher molecular weight (HMW) *n*-alkanoic acid. Numbers mentioned on the peaks refer *n*-alkanoic acid carbon number.

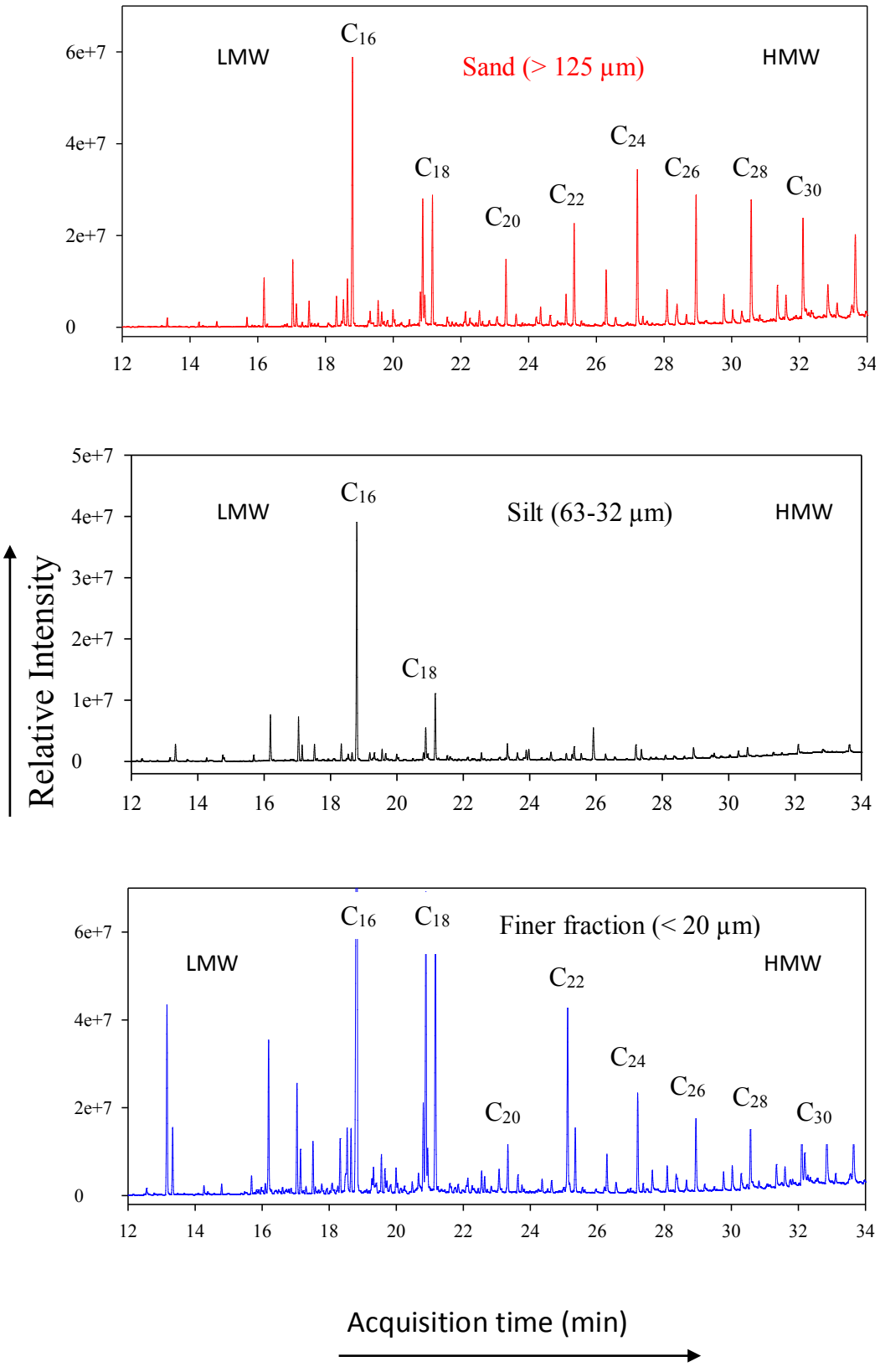


Figure 4d: Representative *n*-alkanoic acid spectra and various chain-lengths of the (a) sand-, (b) silt- and (c) finer size fractions. Numbers mentioned on the peaks refer *n*-alkanoic acid carbon number. LMW- Low molecular weight; HMW- High molecular weight compounds.

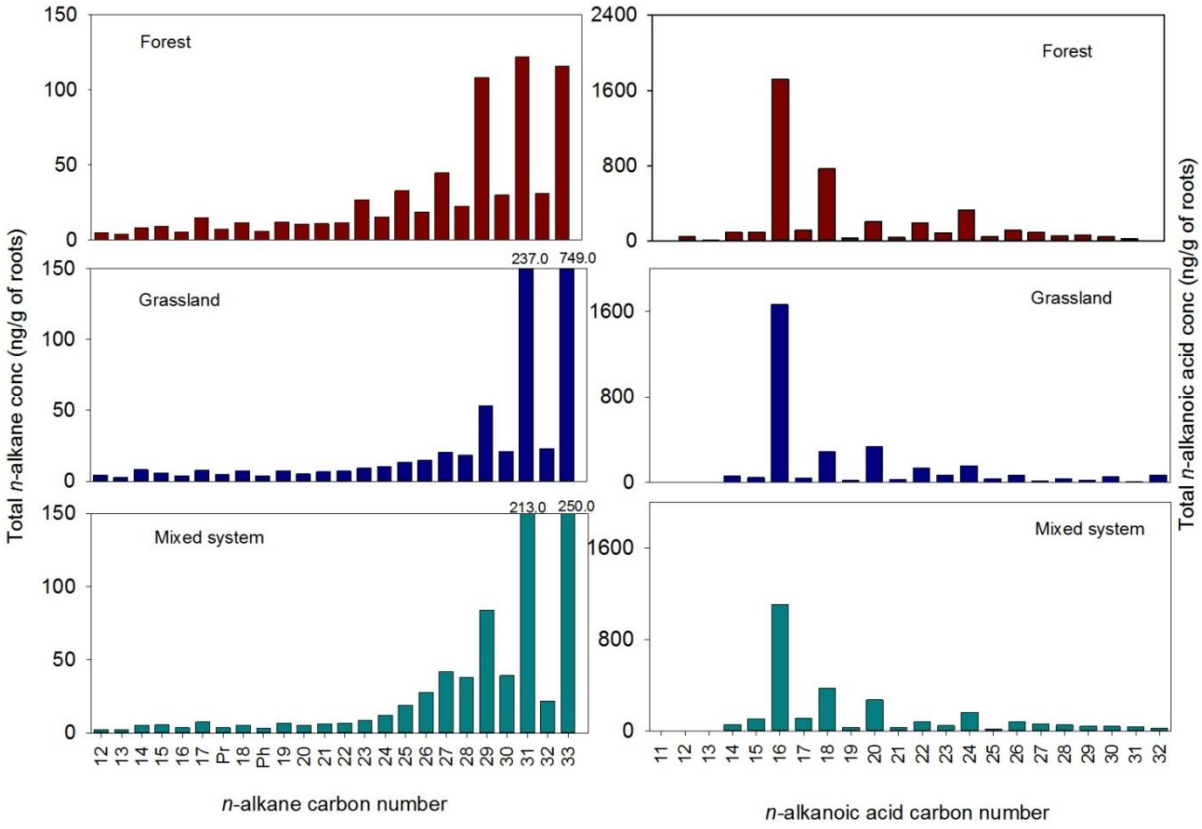


Figure 6a: Abundances of the C_{11}–C_{33} *n*-alkanes and C_{12}–C_{34} *n*-alkanoic acids (in ng *n*-alkyl lipids/g dry roots) and resulting *n*-alkanes (left) and *n*-alkanoic acids (right) distribution patterns in the forest, grassland and mixed system soils (top, middle and bottom, respectively).

6. REFERENCES

- Amblès, A., Jambu, P., Jacquesy, J.-C., Parlanti, E., and Secouet, B.: Changes in the ketone portion of lipidic components during the decomposition of plant debris in a hydromorphic forest-podzol, Soil Science, 156, 49–56, 1993.
- Amblès, A., Jambu, P., Parlanti, E., Joffre, J. and Riffe, C., 1994. Incorporation of natural monoacids from plant residues into an hydromorphic forest podzol. *European Journal of Soil Science*, *45*(2), pp.175-182..
- Anderson, J.A., Hooper, M.J., Zak, J.C. and Cox, S.B., 2009. Characterization of the structural and functional diversity of indigenous soil microbial communities in smelter-impacted and nonimpacted soils. *Environmental Toxicology and Chemistry: An International Journal*, *28*(3), pp.534-541.
- Angst, G., Nierop, K.G., Angst, Š. and Frouz, J., 2018. Abundance of lipids in differently sized aggregates depends on their chemical composition. *Biogeochemistry*, *140*(1), pp.111-125.
- Bergamaschi, B.A., Tsamakis, E., Keil, R.G., Eglinton, T.I., Montluçon, D.B. and Hedges, J.I., 1997. The effect of grain size and surface area on organic matter, lignin and carbohydrate concentration, and molecular compositions in Peru Margin sediments. *Geochimica et Cosmochimica Acta*, *61*(6), pp.1247-1260.
- Bokhorst, S., Huiskes, A., Convey, P. and Aerts, R., 2007. Climate change effects on organic matter decomposition rates in ecosystems from the Maritime Antarctic and Falkland Islands. *Global Change Biology*, *13*(12), pp.2642-2653.
- Brittingham, A., Hren, M. T. and Hartman, G.: Microbial alteration of the hydrogen and carbon isotopic composition of n-alkanes in sediments, Organic Geochemistry, 107, 1–8, https://doi.org/10.1016/j.orggeochem.2017.01.010, 2017.
- Brock, O., Kooijman, A., Nierop, K.G., Muys, B., Vancampenhout, K. and Jansen, B., 2019. Disentangling the effects of parent material and litter input chemistry on molecular soil organic matter composition in converted forests in Western Europe. *Organic Geochemistry*, *134*, pp.66-76.
- Burke, I.C., Yonker, C.M., Parton, W.J., Cole, C.V., Flach, K. and Schimel, D.S., 1989. Texture, climate, and cultivation effects on soil organic matter content in US grassland soils. *Soil science society of America journal*, *53*(3), pp.800-805.
- Carr, A.S., Boom, A., Grimes, H.L., Chase, B.M., Meadows, M.E. and Harris, A., 2014. Leaf wax n-alkane distributions in arid zone South African flora: Environmental controls, chemotaxonomy and palaeoecological implications. *Organic Geochemistry*, *67*, pp.72-84.
- Cayet, C. and Lichtfouse, E., 2001. δ13C of plant-derived n-alkanes in soil particle-size fractions. *Organic Geochemistry*, *32*(2), pp.253-258.

- Chaffee, A.L., Hoover, D.S., Johns, R.B. and Schweighardt, F.K., 1986. Biological markers extractable from coal. In *Biological markers in the sedimentary record* (pp. 311-345). Elsevier.
- Cheng, W. and Coleman, D.C., 1990. Effect of living roots on soil organic matter decomposition. *Soil Biology and Biochemistry*, *22*(6), pp.781-787.
- Chenu, C., Stotzky, G., Huang, P. and Bollag, J., 2002. Interactions between microorganisms and soil particles: an overview. Interactions between soil particles and microorganisms: Impact on the terrestrial ecosystem, 1, pp.1-40.
- Chikaraishi, Y., Naraoka, H. and Poulson, S.R., 2004. Carbon and hydrogen isotopic fractionation during lipid biosynthesis in a higher plant (Cryptomeria japonica). *Phytochemistry*, *65*(3), pp.323-330.
- Clemente, J.S., Simpson, A.J. and Simpson, M.J., 2011. Association of specific organic matter compounds in size fractions of soils under different environmental controls. Organic Geochemistry, 42(10), pp.1169-1180.
- Cranwell, P.A., 1981. Diagenesis of free and bound lipids in terrestrial detritus deposited in a lacustrine sediment. *Organic Geochemistry*, *3*(3), pp.79-89.
- Diefendorf, A.F. and Freimuth, E.J., 2017. Extracting the most from terrestrial plant-derived n-alkyl lipids and their carbon isotopes from the sedimentary record: A review. *Organic Geochemistry*, *103*, pp.1-21.
- Dinel, H., Schnitzer, M., Mehuys, G.R., 1990. Soil lipids: origin, nature, content, decomposition,and effect on soil physical properties. In: Bollag, J.-.M. (Ed.), Soil Biochemistry. Marcel Kekker Inc., New York, pp. 397–429.
- Eglinton, G. and Hamilton, R.J., 1963. The distribution of alkanes. Chemical plant taxonomy, 187, p.217.
- Eglinton, G. and Hamilton, R.J., 1967. Leaf epicuticular waxes. *Science*, *156*(3780), pp.1322-1335.
- Fang, J., Kawamura, K., Ishimura, Y. and Matsumoto, K., 2002. Carbon isotopic composition of fatty acids in the marine aerosols from the western North Pacific: implication for the source and atmospheric transport. *Environmental science & technology*, *36*(12), pp.2598-2604.
- Feeney, D.S., Crawford, J.W., Daniell, T., Hallett, P.D., Nunan, N., Ritz, K., Rivers, M. and Young, I.M., 2006. Three-dimensional microorganization of the soil–root–microbe system. *Microbial ecology*, *52*(1), pp.151-158.
- Feng, X., Simpson, A.J. and Simpson, M.J., 2005. Chemical and mineralogical controls on humic acid sorption to clay mineral surfaces. Organic Geochemistry, 36(11), pp.1553-1566.
- Gérard, F., 2016. Clay minerals, iron/aluminum oxides, and their contribution to phosphate sorption in soils—A myth revisited. *Geoderma*, *262*, pp.213-226.
- Ghosh, S., Wang, Z.-Y., Kang, S., Bhowmik, P.C., Xing, B., 2009. Sorption and fractionation of a peat derived humic acid by kaolinite, montmorillonite and goethite. Pedosphere 19, 21–30.
- Griepentrog, M., Bodé, S., Boeckx, P. and Wiesenberg, G.L., 2016. The fate of plant wax lipids in a model forest ecosystem under elevated CO2 concentration and increased nitrogen deposition. *Organic Geochemistry*, *98*, pp.131-140.

- Griepentrog, M., Eglinton, T.I., Hagedorn, F., Schmidt, M.W. and Wiesenberg, G.L., 2015. Interactive effects of elevated CO2 and nitrogen deposition on fatty acid molecular and isotope composition of above-and belowground tree biomass and forest soil fractions. *Global Change Biology*, *21*(1), pp.473-486.
- Grimalt, J. O., Torras, E. and Albaigés, J.: Bacterial reworking of sedimentary lipids during sample storage, Organic Geochemistry, 13, 741–746, https://doi.org/10.1016/0146-6380(88)90096-4, 1988.
- Hemkemeyer, M., Dohrmann, A.B., Christensen, B.T. and Tebbe, C.C., 2018. Bacterial preferences for specific soil particle size fractions revealed by community analyses. Frontiers in microbiology, 9, p.149.
- Howard, S., McInerney, F.A., Caddy-Retalic, S., Hall, P.A. and Andrae, J.W., 2018. Modelling leaf wax n-alkane inputs to soils along a latitudinal transect across Australia. *Organic Geochemistry*, *121*, pp.126-137.
- Huang, P.M., Wang, M.K. and Chiu, C.Y., 2005. Soil mineral–organic matter–microbe interactions: impacts on biogeochemical processes and biodiversity in soils. Pedobiologia, 49(6), pp.609-635.
- Huang, X., Wang, C., Zhang, J., Wiesenberg, G.L., Zhang, Z. and Xie, S., 2011. Comparison of free lipid compositions between roots and leaves of plants in the Dajiuhu Peatland, central China. *Geochemical Journal*, *45*(5), pp.365-373.
- Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., and Schulze, E. D.: A global analysis of root distributions for terrestrial biomes, Oecologia, 108, 389–411, 1996.
- Jansen, B., Nierop, K.G., Hageman, J.A., Cleef, A.M. and Verstraten, J.M., 2006. The straight-chain lipid biomarker composition of plant species responsible for the dominant biomass production along two altitudinal transects in the Ecuadorian Andes. *Organic Geochemistry*, *37*(11), pp.1514-1536.
- Jansen, B. and Wiesenberg, G.L., 2017. Opportunities and limitations related to the application of plant-derived lipid molecular proxies in soil science. *Soil*, *3*(4), pp.211-234.
- Jones, E. and Singh, B., 2014. Organo-mineral interactions in contrasting soils under natural vegetation. *Frontiers in Environmental Science*, *2*, p.2.
- Kallenbach, C.M., Frey, S.D. and Grandy, A.S., 2016. Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. *Nature communications*, *7*(1), pp.1-10.
- Khatoon, H., Solanki, P., Narayan, M., Tewari, L., Rai, J. and Hina Khatoon, C., 2017. Role of microbes in organic carbon decomposition and maintenance of soil ecosystem. *International Journal of Chemical Studies*, *5*(6), pp.1648-1656.
- Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry*, *34*(2), pp.139-162.
- Kögel-Knabner, I., Guggenberger, G., Kleber, M., Kandeler, E., Kalbitz, K., Scheu, S., Eusterhues, K. and Leinweber, P., 2008. Organo-mineral associations in temperate soils: Integrating biology, mineralogy, and organic matter chemistry. *Journal of Plant Nutrition and Soil Science*, *171*(1), pp.61-82.

- Kolattukudy, P.E., 1996. Biosynthetic pathways of cutin and waxes, and their sensitivity to environmental stress. *Plant cuticles: an integrated functional approach*.
- Kolattukudy, P.E., Croteau, R. and Buckner, J.S., 1976. Biochemistry of plant waxes. *Chemistry and biochemistry of natural waxes*.
- Kuhn, T.K., Krull, E.S., Bowater, A., Grice, K. and Gleixner, G., 2010. The occurrence of short chain n-alkanes with an even over odd predominance in higher plants and soils. *Organic Geochemistry*, *41*(2), pp.88-95.
- Kunst, L. and Samuels, A.L., 2003. Biosynthesis and secretion of plant cuticular wax. *Progress in lipid research*, *42*(1), pp.51-80.
- Lehmann, J. and Kleber, M., 2015. The contentious nature of soil organic matter. *Nature*, *528*(7580), pp.60-68.
- Lichtfouse, E., 1998. Isotope and biosynthetic evidence for the origin of long-chain aliphatic lipids in soils. *Die Naturwissenschaften*, *85*(2), pp.76-77.
- Lichtfouse, E., 1999. Temporal pools of individual organic substances in soil. *Analysis*, *27*(5), pp.442-446.
- Lichtfouse, E., Leblond, C., Da Silva, M. and Béhar, F., 1998. Occurrence of biomarkers and straight-chain biopolymers in humin: implication for the origin of soil organic matter. *Naturwissenschaften*, *85*(10), pp.497-501.
- Lipid abundance and composition of a humic Oxisol as a function of land use Cristiane Pereira de Assis 2010
- Lockheart, M.J., Van Bergen, P.F. and Evershed, R.P., 1997. Variations in the stable carbon isotope compositions of individual lipids from the leaves of modern angiosperms: implications for the study of higher land plant-derived sedimentary organic matter. *Organic Geochemistry*, *26*(1-2), pp.137-153.
- Lützow, M.V., Kögel-Knabner, I., Ekschmitt, K., Matzner, E., Guggenberger, G., Marschner, B. and Flessa, H., 2006. Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions–a review. *European journal of soil science*, *57*(4), pp.426-445.
- Mambelli, S., Bird, J.A., Gleixner, G., Dawson, T.E. and Torn, M.S., 2011. Relative contribution of foliar and fine root pine litter to the molecular composition of soil organic matter after in situ degradation. *Organic Geochemistry*, *42*(9), pp.1099-1108.
- Martelanc, M., Vovk, I. and Simonovska, B., 2007. Determination of three major triterpenoids in epicuticular wax of cabbage (Brassica oleracea L.) by high-performance liquid chromatography with UV and mass spectrometric detection. *Journal of Chromatography A*, *1164*(1-2), pp.145-152.
- Matsuda, H. and Koyama, T., 1977. Early diagenesis of fatty acids in lacustrine sediments—I. Identification and distribution of fatty acids in recent sediment from a freshwater lake. *Geochimica et Cosmochimica Acta*, *41*(6), pp.777-783.
- Meyers, P.A., 2003. Applications of organic geochemistry to paleolimnological reconstructions: a summary of examples from the Laurentian Great Lakes. *Organic geochemistry*, *34*(2), pp.261-289.
- Meyers, P.A. and Ishiwatari, R., 1993. Lacustrine organic geochemistry—an overview of indicators of organic matter sources and diagenesis in lake sediments. *Organic geochemistry*, *20*(7), pp.867-900.

- Mikutta, R., Mikutta, C., Kalbitz, K., Scheel, T., Kaiser, K. and Jahn, R., 2007. Biodegradation of forest floor organic matter bound to minerals via different binding mechanisms. Geochimica et Cosmochimica Acta, 71(10), pp.2569-2590
- Miltner, A., Bombach, P., Schmidt-Brücken, B. and Kästner, M., 2012. SOM genesis: microbial biomass as a significant source. Biogeochemistry, 111(1), pp.41-55.
- Moscatelli, M.C., Secondi, L., Marabottini, R., Papp, R., Stazi, S.R., Mania, E. and Marinari, S., 2018. Assessment of soil microbial functional diversity: land use and soil properties affect CLPP-MicroResp and enzymes responses. *Pedobiologia*, *66*, pp.36-42.
- Moucawi, J., Fustec, E., Jambu, P. and Jacquesy, R., 1981. Decomposition of lipids in soils: free and esterified fatty acids, alcohols and ketones. *Soil Biology and Biochemistry*, *13*(6), pp.461-468..
- Mueller, K.E., Polissar, P.J., Oleksyn, J. and Freeman, K.H., 2012. Differentiating temperate tree species and their organs using lipid biomarkers in leaves, roots and soil. *Organic Geochemistry*, *52*, pp.130-141.
- Otto, A. and Simpson, M.J., 2005. Degradation and preservation of vascular plant-derived biomarkers in grassland and forest soils from Western Canada. Biogeochemistry, 74(3), pp.377-409.
- Paustian, K., Six, J., Elliott, E.T. and Hunt, H.W., 2000. Management options for reducing CO2 emissions from agricultural soils. *Biogeochemistry*, *48*(1), pp.147-163.
- Peters, K.E., Peters, K.E., Walters, C.C. and Moldowan, J.M., 2005. The biomarker guide (Vol. 1). Cambridge university press.
- Prescott, C.E. and Grayston, S.J., 2013. Tree species influence on microbial communities in litter and soil: current knowledge and research needs. *Forest Ecology and Management*, *309*, pp.19-27.
- Preston, C.M., Shipitalo, S.E., Dudley, R.L., Fyfe, C.A., Mathur, S.P. and Levesque, M., 1987. Comparison of 13C CPMAS NMR and chemical techniques for measuring the degree of decomposition in virgin and cultivated peat profiles. Canadian Journal of Soil Science, 67(1), pp.187-198.
- Quenea, K., Derenne, S., Largeau, C., Rumpel, C. and Mariotti, A., 2004. Variation in lipid relative abundance and composition among different particle size fractions of a forest soil. *Organic Geochemistry*, *35*(11-12), pp.1355-1370.
- Quideau, S.A., Chadwick, O.A., Benesi, A., Graham, R.C. and Anderson, M.A., 2001. A direct link between forest vegetation type and soil organic matter composition. *Geoderma*, *104*(1-2), pp.41-60.
- Rodionov, A., Amelung, W., Urusevskaja, I. and Zech, W., 2001. Origin of the enriched labile fraction (ELF) in Russian Chernozems with different site history. Geoderma, 102(3-4), pp.299-315.
- Rommerskirchen, F., Plader, A., Eglinton, G., Chikaraishi, Y. and Rullkötter, J., 2006. Chemotaxonomic significance of distribution and stable carbon isotopic composition of long-chain alkanes and alkane-1-ols in C4 grass waxes. *Organic Geochemistry*, *37*(10), pp.1303-1332.
- Roy, B., Ghosh, S. and Sanyal, P., 2020. Impact of monsoon, vegetation, and landscape on pedogenesis: A case study using organic and inorganic tracers from the

Himalayan foreland sediments. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *556*, p.109854.

- Rustan, A.C. and Drevon, C.A., 2001. Fatty acids: structures and properties. e LS.
- Schäfer, I.K., Lanny, V., Franke, J., Eglinton, T.I., Zech, M., Vysloužilová, B. and Zech, R., 2016. Leaf waxes in litter and topsoils along a European transect. *Soil*, *2*(4), pp.551-564.
- Schefuß, E., Ratmeyer, V., Stuut, J.B.W., Jansen, J.F. and Damsté, J.S.S., 2003. Carbon isotope analyses of n-alkanes in dust from the lower atmosphere over the central eastern Atlantic. *Geochimica et Cosmochimica Acta*, *67*(10), pp.1757-1767.
- Schmidt, M.W., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A. and Nannipieri, P., 2011. Persistence of soil organic matter as an ecosystem property. *Nature*, *478*(7367), pp.49-56.
- Selesi, D., Pattis, I., Schmid, M., Kandeler, E. and Hartmann, A., 2007. Quantification of bacterial RubisCO genes in soils by cbbL targeted real-time PCR. *Journal of microbiological methods*, *69*(3), pp.497-503.
- Sessitsch, A., Weilharter, A., Gerzabek, M.H., Kirchmann, H. and Kandeler, E., 2001. Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. Applied and environmental microbiology, 67(9), pp.4215-4224.
- Shiau, Y.J., Chen, J.S., Chung, T.L., Tian, G. and Chiu, C.Y., 2017. 13 C NMR spectroscopy characterization of particle-size fractionated soil organic carbon in subalpine forest and grassland ecosystems. *Botanical studies*, *58*(1), pp.1-7.
- Six, J., Bossuyt, H., Degryze, S. and Denef, K., 2004. A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil and Tillage Research, 79(1), pp.7-31.
- Six, J. and Paustian, K., 2014. Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biology and Biochemistry*, *68*, pp.A4-A9.
- Soil physical attributes author links open overlay panel Daniel Hillel (https://doi.org/10.1016/B978-0-12-348536-6.50010-1)
- Sokol, N.W., Sanderman, J. and Bradford, M.A., 2019. Pathways of mineral-associated soil organic matter formation: Integrating the role of plant carbon source, chemistry, and point of entry. *Global Change Biology*, *25*(1), pp.12-24.
- Sollins, P., Homann, P. and Caldwell, B.A., 1996. Stabilization and destabilization of soil organic matter: mechanisms and controls. *Geoderma*, *74*(1-2), pp.65-105.
- Spohn, M., 2018. Phosphorus and carbon in soil particle size fractions–A global synthesis. *Biogeosciences Discussions*, pp.1-24.
- Sposito, G., Skipper, N.T., Sutton, R., Park, S.H., Soper, A.K. and Greathouse, J.A., 1999. Surface geochemistry of the clay minerals. *Proceedings of the National Academy of Sciences*, *96*(7), pp.3358-3364.
- Thompson, S. and Eglinton, G., 1978. The fractionation of a recent sediment for organic geochemical analysis. *Geochimica et Cosmochimica Acta*, *42*(2), pp.199-207.

- Wardle, D.A., Bardgett, R.D., Klironomos, J.N., Setälä, H., Van Der Putten, W.H. and Wall, D.H., 2004. Ecological linkages between aboveground and belowground biota. *Science*, *304*(5677), pp.1629-1633.
- Weng, X., Li, J., Sui, X., Li, M., Yin, W., Ma, W., Yang, L. and Mu, L., 2021. Soil microbial functional diversity responses to different vegetation types in the Heilongjiang Zhongyangzhan Black-billed Capercaillie Nature Reserve. *Annals of Microbiology*, *71*(1), pp.1-11.
- Wiesenberg, G.L., Dorodnikov, M. and Kuzyakov, Y., 2010. Source determination of lipids in bulk soil and soil density fractions after four years of wheat cropping. *Geoderma*, *156*(3-4), pp.267-277..
- Wiesenberg, G.L., Schmidt, M.W. and Schwark, L., 2008. Plant and soil lipid modifications under elevated atmospheric CO2 conditions: I. Lipid distribution patterns. *Organic Geochemistry*, *39*(1), pp.91-102.
- Wiesenberg, G.L., Schwark, L. and Schmidt, M.W., 2004. Improved automated extraction and separation procedure for soil lipid analyses. *European Journal of Soil Science*, *55*(2), pp.349-356.
- Wiesenberg, G.L., Schwarzbauer, J., Schmidt, M.W. and Schwark, L., 2004. Source and turnover of organic matter in agricultural soils derived from n-alkane/n-carboxylic acid compositions and C-isotope signatures. *Organic Geochemistry*, *35*(11-12), pp.1371-1393.
- Wiesenberg, G.L. and Schwark, L., 2006. Carboxylic acid distribution patterns of temperate C3 and C4 crops. *Organic geochemistry*, *37*(12), pp.1973-1982.
- Yakimovich, K.M., Emilson, E.J., Carson, M.A., Tanentzap, A.J., Basiliko, N. and Mykytczuk, N., 2018. Plant litter type dictates microbial communities responsible for greenhouse gas production in amended lake sediments. *Frontiers in microbiology*, *9*, p.2662.
- Yang, S., Jansen, B., Absalah, S., Kalbitz, K. and Cammeraat, E.L., 2020. Selective stabilization of soil fatty acids related to their carbon chain length and presence of double bonds in the Peruvian Andes. Geoderma, 373, p.114414.
- Young, I. M., Crawford, J. W., Nunan, N., Otten, W. & Spiers, A. Microbial distribution in soils: physics and scaling. Adv. Agron. 100, 81–121 (2008).