Designing a Soil Health Index for Sustainable Agricultural Systems

Munisath Jahan Khandoker

Lancaster Environment Centre, Lancaster University

This thesis is submitted for the degree of
Doctor of Philosophy

March 2024

ROTHAMSTED
RESEARCH

Lancaster
Environment Centre

Lancaster
University

Designing a Soil Health Index for Sustainable Agricultural Systems.

Munisath J Khandoker

Thesis Abstract

Soil health is vital for agriculture and agro-ecosystems. Healthy soils act as a reservoir and cycling system for water, air, organic matter, and nutrients essential for crop growth and as a habitat for a diverse array of organisms, including bacteria, fungi, insects, and worms, contributing towards ecosystem stability and resilience. However, soil health cannot be directly measured effectively with one indicator. Instead, soil health assessments typically rely on a range of measurements of essential biological, physical, and chemical indicators. Due to the complexity and highly integrative nature of soils, it is difficult to develop general soil health indices.

The main goal of this thesis was to develop a soil health index that was able to quantify soil health for different agricultural land uses and soil textures across the UK. With the aid of Rothamsted Research's long-term experiments of known history and land management, we aimed to collect measurements of crucial physical, biological, and chemical soil health indicators using traditional methods. Then, using structural equation modelling (SEM), we hoped to create a robust soil health assessment. When designing a soil health index, selecting the most informative indicators of soil health is essential. To achieve this, we initially focused on enhancing existing soil monitoring methodologies and exploring novel technologies that can streamline this process. In Chapter 2, we evaluated the SLAKES smartphone application as a tool for measuring soil stability. Results from the SLAKES application were compared with the established Le Bissonnais method. Results showed SLAKES could differentiate between different management types on clayey soil, but were less sensitive when tested with sandy soil. Despite this lower sensitivity, we conclude that the SLAKES app can be a legitimate method to measure aggregate stability, providing a faster and easier method for researchers and land managers compared to conventional methods.

In Chapter 3, we evaluated the use of extracellular enzymes, N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU), as promising soil health indicators. The objective was to investigate which of these soil enzymes, if any, could be used as a comprehensive biological indicator for soil health by examining the relationships between microbial enzyme activity in a range of soils with contrasting chemical and physical properties. We observed that grass treatments relative to all other plots showed increased levels of enzymatic activity, followed by arable and fallow, respectively. Furthermore, enzyme activity correlated with other observed soil health indicators.

Using learnings from the previous two chapters, the Chapter 4 study aimed to create meaningful metrics for soil health that influence agricultural production and other ecosystem services. We included a range of soil measurements relevant to soil health, including physical, chemical, and biological soil indicators under contrasting agricultural land uses and soil types. We found that SEM allows for a comprehensive understanding of the complex relationships within the soil health system. It was particularly beneficial when dealing with multiple indicators and latent variables that contribute to overall soil health. Based on our results, we believe soil scientists can leverage SEM to refine soil health assessment models and improve the accuracy of their measurements, as well as understand the effects of agricultural management practices on soil health.

From selecting the most informative soil health indicators to carefully considering the measurement methods used for these indicators, this study demonstrates the multifaceted nature of designing a soil health index.

Table of Contents

Author's Declaration

I declare that this thesis is my own work and has not been submitted in substantially the same form for the award of a higher degree elsewhere. This thesis has been prepared as a set of papers intended for submission to peer-reviewed journals. Chapters 2, 3, and 4 are structured in accordance with the format intended for journal submission. A combined reference list can be found at the end of the thesis.

Chapter 1 provides a general introduction to the research area and the aims, objectives, and key hypotheses of the thesis. It is not intended for publication.

Chapter 2 has been submitted for publication.
Khandoker, M., Haefele, S., Gregory, A., Ostle, N. (2023).
MK designed the research and carried out the data collection and analysis with help from SH and AG. MK prepared the manuscript with input from SH, AG, and NO.
Chapter 2 has been submitted for publication in the European Journal of Soil Science and has undergone a first review.

Chapter 3 is intended for publication.
Khandoker, M., Haefele, S., Gregory, A., Ostle, N. (2023).
MK designed the research and carried out the data collection and analysis with help from SH and AG. MK conducted laboratory work and analysed results. MK prepared the manuscript with input from SH, AG, and NO.

Chapter 4 is intended for publication.
Khandoker, M., Haefele, S., Gregory, A., Metcalfe, H., Ostle, N. (2023).
MK and HM designed the research. MK, with the help of SH and AG, conducted sampling. MK conducted laboratory work and analysed results. MK prepared the manuscript with input from SH, AG, and HM.

Chapter 5 comprises a general discussion and conclusions and is not intended for publication.

Acknowledgements

This thesis would not have been possible without the contribution and support of many colleagues and friends. First, I would like to thank my supervisors, Dr Stephan Haefele, Dr Andy Gregory, and Professor Nick Ostle, for their wisdom over the last few years. A special thank you to Stephan and Andy for their help during fieldwork on the hottest and coldest days of the year! I am profoundly grateful to both of you for your unwavering support and guidance throughout my PhD journey. Your expertise, mentorship, and encouragement have played a pivotal role in defining my academic development and personal growth.

Rothamsted Research has been an amazing workplace. Thanks to Rothamsted's statistical team and the Sustainable Soils group. Being a part of this research group has been immensely rewarding, with everyone readily offering guidance and support throughout my doctoral studies. I would also like to express my sincere thanks to Helen Metcalfe for her invaluable support and expertise throughout my research journey.

Finally, I'd like to thank my family and friends for their support, the constant supply of tea, for putting up with the constant thesis chat over the last months, and for pretending to be interested in soil. A special thanks to Nishath for her invaluable help during the final months of my PhD journey, my parents for their sacrifices, patience, and belief in my abilities, which have fuelled my determination to reach this milestone, to my husband, Shakir, for tolerating my enthusiasm for all things soil and recognising the value of my work, and to Asiyah and Sharmin for bringing constant joy to my life.

All praise and thanks to God almighty.

List of Figures

List of Tables

1 Introduction

A typical topsoil is composed of around 45% minerals, 5% organic matter, and 50% water and air. However, these percentages are broad approximations as soils are highly intricate and dynamic, and their composition can vary daily, influenced by factors like water availability, land management, and soil type (Spork et al., 2002). This layer, supported by the subsoil below, serves as a foundational element for our food system. It holds essential nutrients necessary for plant growth, functions as a water reservoir, and supports diverse ecosystems. As a natural water filter, soil plays a crucial role in providing drinking water and serves as a foundation for construction. Additionally, it serves as the Earth's primary terrestrial carbon (C) storage system, contributing significantly to climate regulation (Jónsson et al., 2016). Carbon stored in soil worldwide (1200–2400 Pg) exceeds that stored in the atmosphere (720–750 Pg) and terrestrial plants (550–835 Pg) combined (Batjes. 1996; Scharlemann et al., 2014).

Despite its vital role in sustaining life, soil is facing a crisis globally. Intensive agriculture, coupled with the impact of a growing population, has led to significant shifts in land use and soil contamination. Remarkably, it is estimated that only 3 cm of topsoil is generated every 1,000 years (UN FAO, 2019); however, in England and Wales, an alarming 2.9 million tonnes of topsoil are eroded annually. According to the Environment Agency (2019), the majority of soils in England fall under the classifications of either 'degraded' or 'very degraded.' Land degradation has far-reaching consequences, affecting food security, water availability, and ecosystem health, directly impacting half of the global population. This degradation results in an annual loss of approximately US$40 trillion worth of ecosystem services, which accounts for nearly half of the global GDP in 2021 (United Nations, 2022).

Furthermore, once the soil reaches a degraded state, it has the potential to emit greenhouse gases, such as carbon dioxide, methane, and nitrous oxide, and suffer from loss of organic matter, contamination, compaction, increased salinity, and other detrimental effects (UN FAO, 2019; European Commission, 2023). This degradation can be abrupt, but oftentimes, degradation is a more gradual process, impacting agricultural production and the broader

environment over time. In response, research has focused on developing measures to assess soil health, aiming to monitor its condition and guide management practices to prevent degradation. This has sparked debates around the fundamental question: “What is soil health?”

In the agricultural context, soil health can be defined as a soil's ability to respond to agricultural interventions. This definition aligns closely with 'soil quality,' which pertains to a soil's condition and properties relative to the requirements of one or more species, including humans (Kibblewhite et al., 2007). Soil quality also underscores the dynamic nature of soil, reflecting properties influenced by land management practices. However, these definitions might oversimplify soil health by focusing solely on its benefits for agricultural productivity and neglect the temporal and geographical scales over which soil influences entire ecosystems. It also fails to acknowledge the integrated nature of the biological, physical, and chemical qualities inherent in soil. Thus, a better understanding of soil components and their connections with each other would lead to a more holistic approach to characterising soil functioning; such approaches have been used to define the health of soils.

1.1 Historical Perspectives of Soil Health

Across history, humans have formed perceptions of soil functionality, which are evident as far back as 2070 BCE when ancient Chinese texts discussed assessing soil suitability for crop growth (Bünemann et al., 2018). It is reasonable to assume that even our hunter-gatherer ancestors recognised variations in soil productivity, noting that certain areas were more adept at producing certain plants (Feeney, 2019). As societies transitioned to settled living, villages and later cities tended to emerge in areas where land could sustain the population's food needs. The recent expansive growth of cities has adversely affected global food security, as particularly productive land is lost. Hillel (1992) contends that the decline of some ancient civilisations can be linked to their inability to preserve essential soil properties and functions vital for food production.

The term "soil fertility" (SF) has traditionally referred to the ability of specific soils to aid the growth of crops (Patzel et al., 2009). However, the broader ecosystem functions of soils and the importance of non-agricultural lands, such as forests or savannahs, were not well recognised during this time. Consequently, this term has inherent limitations in capturing the full spectrum of soil's roles and functions. Even within agriculture, soil fertility was commonly interpreted as focusing exclusively on the availability of crop nutrients, possibly overlooking other crucial soil properties.

"Soil quality" (SQ) is commonly attributed to a 1977 conference paper by Warkentin and Fletcher. However, Bünemann et al. (2018) reference an earlier work by Mausel (1971), which discussed the suitability of different soils for growing crops in Illinois, USA. Here, Mausel (1971) describes SQ as the capacity of soils to produce crops such as corn, soybeans, and wheat under conditions of intensive management, where the selection of crops as an indicator of soil quality is attributed to their significant economic dominance in the agricultural sector.

During the 1980s and 1990s, the term experienced a significant surge in popularity and widespread acceptance. During this period, a conceptual framework for the term was developed, reflecting the growing awareness among land managers about the crucial role of soil in delivering ecosystem services. Another perceived advantage of the term SQ was its accessibility to non-specialists. The belief was that the concept could be more easily grasped by a broader audience, including politicians and policymakers, compared to a more technical description of specific soil properties or functions. This was particularly relevant as the term "quality" was increasingly employed to describe the attributes of water and air.

Doran et al. (1994) noted that the current concept of SQ was limited due to its emphasis on agriculture and production. They advocated for a definition of SQ that highlighted the primary concerns associated with soil use. They incorporated into their definition that in addition to productivity, soils have the capacity to contribute towards environmental quality and promote plant, animal, and even human health (Bunemann et al., 2018).

The concept of a healthy soil initially emerged during the early twentieth century in Europe when organic farming rose in popularity amongst farmers. However, in the last 20 years, we have seen the widespread use of the term “soil health” (SH) gradually overtaking SQ (Figure 1). There is debate as to whether the terms SH and SQ are synonymous, which Bünermann et al. (2018) discuss in some detail. Some scientists argue that SH captures the biological aspects of the soil better and is a more comprehensive term. However, this might be just a matter of definition, as some soil quality assessments incorporate biological properties into their framework.

The 25-Year Environment Plan of the UK government, as outlined by DEFRA (2018), explicitly addresses SH. It contains sections dedicated to "Improving soil health" and "Developing better information on soil health", as well as a commitment to formulate a Soil Health Index (SHI). Similarly, the Government of India introduced Soil Health Management as part of the National Mission for Sustainable Agriculture in 2014 (Government of India, 2023). Here, farmers are provided with laboratory analyses and soil health cards, with funding from the government on national or regional levels. Some tests include pH, electrical conductivity, organic C, and plant-available nutrients. While this set of analyses would have been referred to as an SF assessment fifty years ago, a deliberate decision was made to use the term SH for effective communication with farmers. Similarly, in the UK, the agricultural consultancy group ADAS employs the term SH to communicate its soil management activities with farmers.

There are two main ways to approach SH (Kibblewhite et al., 2007). The reductionist method estimates the condition of the soil using physical, chemical, and biological properties of the soil. In contrast, the integrated approach assumes that SH is more complex than just the soil’s physical, chemical, and biological components and allows for emergent properties resulting from interactions of these processes and properties in the soil (Kibblewhite et al., 2007). Our working definition of SH is derived from a combination of approaches, assuming that soil health in relation to agricultural soils is a state that can support sufficient food and fibre production whilst delivering essential ecosystem services.

Whilst the concept of SH has become widely accepted, a new debate arose as to whether SH could be quantified in a meaningful way. Karlen et al. (1997) were some of the first to promote the quantification of SH. This was followed by significant opposition, namely by Sojka et al. (1999), who argued that SH should not be quantified into a numerical value. Whilst some thought that a soil health index (SHI) was a useful way to summarise data, others claimed that informed management decisions primarily rely on only a few soil properties. As a result, an SHI would generalise information and would not allow land managers to understand the cause of a low SH. For example, fertiliser inputs would need to be applied to correct for low nutrient levels. This would not be readily known if SH was generalised into a single numerical number.

Despite this, several attempts were made to quantify SH into a single value. Stockdale et al. (2003) combined various soil properties but found that doing so would be counterproductive. They found that soil use changes our working definition of SH. For example, soils intended for cultivating horticultural crops typically require a neutral or alkaline pH and high nutrient concentrations. In comparison, soils chosen for growing coniferous trees generally need an acidic pH and lower nutrient levels (Shi et al., 2008). Soil scientists should, therefore, exercise caution against over-complicating the subject of SH. The term SH is meaningless unless it is linked to a function or specified use. Especially when making informed land management decisions, measurements of specific soil properties are of paramount importance.

Figure 1: Changes in the perspectives of soil health over time.

In summary, SH may be defined as “the capacity of a specific kind of soil to function, within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and support human health and habitation” (Doran et al., 2000). SH can be affected by soil use and management approaches (Bunemann

et al., 2018). Thus, knowledge of the effects of different soil management practices and land uses on soil functions is needed in order to develop a working SHI that accounts for these complex interactions (Morrow et al., 2016).

1.2 Ecosystem Services

While the function of soil was recognised long ago, the importance of conservation and enhancement of ecosystem services (ES) rendered by the soil has only been realised more recently. ES refers to the various benefits soils provide to ecosystems and human well-being (Pereira et al., 2018). These services are essential for maintaining ecological functions, supporting biodiversity, sustaining agricultural productivity, and natural resources. There are currently several frameworks that classify ES, such as the Common Classification of Ecosystem Services (CICES). The CICES provide a standardised and consistent way of describing ES. It offers a systematic approach to understanding the different ways ecosystems contribute towards human well-being.

CICES was developed by the European Environment Agency (EEA) and has been widely adopted in environmental and ecological research. The hierarchical classification organises ES into main classes, divisions, groups, and classes (CICES, 2024; Maes et al., 2020). The main classes include services that are involved in the provision of products directly from ecosystems, such as food and water, services that regulate and maintain environmental processes, for example, disease and climate regulation, and supporting services that are necessary for other ES, including soil formation and nutrient cycling. The hierarchical structure of CICES allows for a detailed and standardised classification of ES, providing a common language for researchers, policymakers, and practitioners. This framework helps in systematically analysing and communicating the contributions of ecosystems and has been applied in various studies and assessments to assess the state of ecosystems, monitor changes, and inform sustainable management practices.

Climate change and human activity are placing increasing pressure on our soils and their functions to maintain and regulate ES. As a result, when soils become degraded, so do ES, threatening essential resources of future generations. Moreover, land management has

significant consequences on the capacity and quality of ES provided. The provision, regulation, and value of soil ES are especially threatened in intensely managed lands. Thus, sustainable practices are crucial in restoring soil ES.

1.3 Factors Affecting Soil Health and Ecosystem Services

Soils are formed through a series of complex processes. The parent material forms the basis of the soil, which weathers to smaller particles, releasing essential plant nutrients in the process. Minerals within the parent material combine with organic matter (OM) in the surface layers, shaping the physical and biotic structure and functioning of the soil (Nordin, 2020). The combination of all these factors leads to the formation of soil profiles with distinct horizons. The layering reflects the numerous processes and interactions that have formed the soil over time and is used to define the soil type.

Soil type influences the physical, chemical, and biological properties of soil that collectively determine the soil's ability to support plant growth and sustain ecosystems. Typically, soil type is determined by the proportions of sand, silt, and clay particles, known as soil texture. Clay refers to particles of <2 µm in diameter, silt particles are typically 2 to 50-63 µm, and sand particles are 50-63 µm to 2 mm in diameter (the threshold diameter between 'sand' and 'silt' differs between different classification schemes) (Breemen et al., 2002). The differences in soil texture influence several factors that contribute towards soil health. For example, sandy soils tend to have larger particles and may have good drainage but lower water and nutrient retention. In contrast, clayey soils have smaller particles, providing higher nutrient retention but potentially poor drainage. Loamy soils, with a balanced mixture of sand, silt, and clay, are often considered ideal for healthy plant development (Tahir et al., 2016). Moreover, different soil types may harbour distinct microbial communities, which play a crucial role in nutrient cycling, organic matter (OM) decomposition, and overall soil fertility (Zhang et al., 2023). Soil type, therefore, plays a vital role in determining several properties that collectively define soil health. Farmers and land managers must consider soil type when implementing practices to enhance soil fertility, prevent erosion, and sustain healthy ecosystems.

A combination of climate and fixed properties, such as soil texture and stone content, support specific conditions for soil habitats. Variable abiotic factors, such as compaction and pH, then influence the state of these habitats. Biotic factors, including C substrate availability and microbial biomass, interact with fixed and variable abiotic properties to establish the overall condition and health of the soil ecosystem (Kibblewhite et al., 2008). Previous land management can significantly alter the health of our soils. For example, salinisation due to poor irrigation systems and the loss of SOM from monoculture cropping all contribute to the degradation of surface horizons. Agricultural practices and land management can alter the soil's physical, chemical, and biological properties, thus controlling soil health. Not all such practices are harmful to all aspects of soil health. For example, the application of fertiliser can increase soil fertility, whilst cultivation can create a seedbed optimal for seed germination.

SOM refers to the organic materials present in the soil, primarily originating from the decomposition of plant and animal residues as well as microbial biomass, while soil organic C (SOC) specifically refers to the C component of SOM (Blanco-Canqui et al., 2013). As a generalisation, approximately 58% of SOM is SOC (the van Bemmelen factor), but this can vary greatly depending on the source of the SOM (Batjes et al., 1996). SOC drives soil systems for their integral role in transferring energy. SOC reacts with clay and minerals to form various organic complexes and can form and stabilise soil aggregates due to its high surface area and charge density (Lal, 2016). Several mechanisms stabilise and protect SOC, including the physical protection mechanisms of adsorption to surfaces or occlusion within aggregates and biochemical composition, which can affect the ease with which SOM can be decomposed. For example, the formation of microaggregates captures SOC and can store SOC deep in the soil away from environmental and human interference for millennia (Six et al., 2000). Aromatic and double-bond hydrocarbons can also coat stable aggregates, providing another layer of protection. Furthermore, microorganisms are involved in SOC transfer and can release substances that inhibit decomposer organisms and enzymes that would otherwise use SOC (Dungait et al., 2012).

Current SOC models predict C pools and C turnover rates. For instance, rapid C turnover provides immediate sources of energy, whilst slower rates indicate lasting C energy

reservoirs that can support soil systems and provide structural soil stability over time (Lal, 2004). Thus, the abundance of different OM fractions may indicate the state of the local soil ecosystem and its functions. Moreover, the dynamics of SOC also impact atmospheric chemistry and C-cycling (Scharlemann et al., 2014). Increasing global SOC to deeper soil horizons would cause a significant drawdown of atmospheric CO_2 below ground. Therefore, it is vital to understand the intricate functions and role of SOC, especially in mitigating global climate change, providing ES, and supporting food security and soil health. We should also note that SOC models are not always based on measurable C pools, nor can they be explicitly used to describe the ecological functioning of soils. As a result, the use of SOC models to assess soil health is restricted.

After SOC, nitrogen (N) and phosphorus (P) cycling affect soil system dynamics and delivery of ES. While many argue that C substrates are the primary limiting factor for microbial activity, increasing evidence suggests that microbes are frequently restricted by N availability instead (Schimel, 2005). Thus, in N-stress conditions, the functional capacity of soil ecosystems is strongly impacted by N supply. The same is true for P, which often limits plant growth if not supplied and is particularly important for biological N fixation in free-living and symbiotic organisms.

Agriculture has been manipulating nutrient supplies through fertiliser additions for centuries. Crops require large amounts of N for crop growth, development, and yield. Thus, farmers regularly apply N fertilisers to fulfil the demand for crop production (Anas et al., 2020). In natural, undisturbed soils, there are typically low atmospheric N inputs, and small N losses due to leaching and emissions are observed as microorganisms quickly assimilate mineral N due to high N demand. In contrast, we see a significant increase in N losses through leaching and the atmosphere in disturbed soils and due to greater OM decomposition rates. Soil health can decline as a result of this, as N-dependent organisms that are vital in supporting soil functions and crop growth are restricted due to reduced N availability in the soil. Sustainable agricultural practices propose to counter nutrient losses during crop production using organic amendments (OA) in order to restore and sustain soil health. Despite this, our understanding of the impact of nutrient additions on the condition of soil biota and soil functioning is limited.

1.4 Assessing Soil Health

Assessing soil health across different soil types, climatic zones, and agricultural systems offers many scientific and policy challenges. Soils are multifaceted systems; thus, complex debate arises when discussing appropriate methods to assess soil health. A single indicator cannot describe all properties of soil health, and it would not be realistic to measure every potential soil health indicator.

Proposals for soil tests are often linked with legal frameworks to protect our soils at national and even international scales (DEFRA, 2020). The Environment Agency (2019), in their 'State of the Environment' review, states that 2 million hectares of soil are threatened by erosion in England and Wales. According to figures from 2020, the consequences of soil compaction and erosion are costing the UK economy over £1.2 billion per year, increasing flood risk and threatening biodiversity and soil fertility (DEFRA, 2023).

Furthermore, the European Commission has authorised a new Soil Monitoring Law, in line with the EU Zero Pollution ambition; this would be the first EU legislation on soils (European Commission, 2023). The ultimate goal is to attain healthy EU soils by 2050 and address key soil risks such as erosion, loss of SOM, contamination, compaction and soil biodiversity losses. The proposal offers a comprehensive soil monitoring framework and a consistent soil health definition. These initiatives provide vital knowledge on the current state of the soil. However, a further consolidative framework is needed to identify soil health indicators and relate previous data to soil health.

There are many practical issues when assessing soil health, as soils are multifunctional and can deliver many integrated functions and services. Thus, a full evaluation of soil status demands extensive testing. Moreover, soils constantly interact with external factors, such as temperature and rainfall, and these interactions are variable and hard to control. Hence, assessments of soil systems need long-term monitoring or multiple readings over time as soils do not respond instantly to environmental changes.

Laboratory tests provide helpful information on soil responses within controlled experimental conditions; however, they are not fully indicative of the natural soil ecosystem performance. In contrast, field assessments may require complex, long-term experiments that cannot be achieved on every land. Whole system assessments are often too expensive or virtually impossible to conduct. As a result, using laboratory tests and field assessments are not always feasible or representative of the soil environment. So, while diagnostic tests or testing certain parameters of the soil can provide enough information to describe soil health, it is important to acknowledge that such assessments may not fully capture the complexity of real soil conditions.

Crucial soil functions, such as C transformations, nutrient cycling, soil structure maintenance, and regulation of microbial populations, all have existing methods for assessing their performance. These are based on specific processes linked to these functions and reveal the current activity that supports ES. Individual functions are related within a network of interactions; thus, soil health tests require a more integrated approach. Soil health assessments can be instead measured by a series of diagnostic tests, separated into chemical, biological and physical measurements.

Soil assessments also need to incorporate soil types within their assessments. Such an approach would be more representative of whole soil systems across particular soil types. From here, we can calculate ranges and trends within each system and see if there are any grounds for comparison between different soil types. Overall, analysing trends and changes in soil conditions allows us to weigh the impact of farming practices and climate change on soil health and ES. However, it is crucial to include soil health indicators that are comprehensive and can describe systems effectively in these assessments.

1.5 Soil Health Indicators

Soils play a crucial role in fulfilling multiple functions simultaneously. These functions are complex and interdependent, making soil health and sustainability vital considerations (McBratney et al., 2014). While the quantification of soil functions is currently an active area of research, it presents a significant challenge. Although machines and sensors can make

predictions, they are unable to measure soil functions directly due to the soils elusive nature (Marchant. 2021). Rather, these functions are viewed as complex properties that arise from the intricate interplay of physical, chemical, and biological processes within the soil. Therefore, when evaluating soil functions, we must rely on measurable soil properties (Vogel et al., 2018) representing a wide range of soil functions. A quantitative assessment of our soils is crucial for informing policies and ensuring their continued productivity. It can also be used to create models that predict the effects of external influences from agriculture and climate change. Addressing this challenge is currently one of the most critical tasks in the field of soil science (Vogel et al., 2019).

Soil health depends on different physical, chemical, and biological attributes of the soil that are important for ecosystem services. Concepts for understanding the chemical and physical properties of soils have been largely accepted by the agricultural community. With significant progress in assessing and managing biological functioning in soils, we are better able to choose the relevant indicators to assess soil health.

Chemical indicators of soil health play a major role in providing nutrients for plants. Some of the most commonly used chemical attributes in soil health assessments are soil pH, cation exchange capacity (CEC), and SOM (Kelly et al., 2009). These are also correlated with crop yields and thus can be easily interpreted and measured. These chemical attributes are also involved in maintaining SOM, plant biomass, and nutrient cycling (Smith et al., 2000).

Idowu et al. (2008) selected 39 soil health parameters and correlated them with crop growth and yield in soils under different soil management practices in the US. They concluded that the most important chemical properties describing soil health are soil pH, phosphorus (P), potassium (K), copper (Cu), zinc (Zn), iron (Fe) and manganese (Mn). Soil pH influences the availability of essential nutrients and correlates directly with plant-available nutrients and microbial activity. P is a critical nutrient for plant growth, involved in energy transfer and various metabolic processes, while K is involved in enzyme activation, water regulation, and plant stress resistance. Cu participates in various biochemical processes necessary for plant growth and development, and Zn is a micronutrient required for enzyme activity and essential for various metabolic processes in plants, ensuring soil fertility.

Furthermore, Fe is essential for chlorophyll synthesis and electron transport in plants, while Mn is involved in photosynthesis, N metabolism, and enzyme activation (Osman. 2012).

Maintaining optimal nutrient levels supports plant growth, development, and health and contributes to the overall resilience of plants against environmental stress (Jouran et al., 2017). While these chemical properties provide valuable information about soil fertility and potential nutrient imbalances, regional variations and specific crop requirements may influence the choice of chemical soil health indicators.

As described previously, SOC concentration is another important attribute of soil as it affects important functional processes, such as N storage, water holding capacity, and aggregate stability, and it also affects microbial activity (Bennet et al., 2010; Silva et al., 2007). SOC largely contributes towards soil fertility; however, SOC can also have some negative effects on soil health. For example, high SOC can reduce pesticide efficiency, thus increasing the frequency of applications. This occurs when soluble SOC aggregates together to form different organic fractions that facilitate the transport of pesticides through the soil (Spark et al., 2002; Sojka et al., 1999).

N is key to plant growth and is considered the most limiting plant nutrient in most natural systems (Lines-Kelly et al., 1998; Cantarella, 2007). Mineral N, such as nitrate or ammonium, organic N and potentially mineralisable N in SOM, are all different ways to store soil N. N-availability in the soil changes rapidly when influenced by the weather, physical soil conditions, microbial activity, and the availability of SOM. Furthermore, N availability can differ greatly depending on the time when samples are taken (Hu et al., 2014). Thus, soil health assessments use models in collaboration with soil tests to analyse the real availability of soil N (Cantarella, 2007).

Like N, P is also a key nutrient for crops and is widely assessed in soil health tests. P helps by stimulating root and plant growth (Lines-Kelly et al., 1998). Soil P is generally available as orthophosphates. However, microbial P and organic P can also become readily available in the soil (Zhang et al., 2006).

SOC, pH and plant-available nutrients are the most common chemical parameters currently being used in soil health assessments (Jurandy et al., 2013). However, soil health tests that include chemical indicators alone cannot provide a representative view of soil health. Instead, the integration of chemical, biological and physical indicators is the most suitable approach when assessing soil health.

Measuring the physical attributes of soil largely consists of simple, low-cost methods. Furthermore, physical indicators of soil health, such as aggregate stability, bulk density, and porosity, can be correlated to hydrological processes such as erosion, water holding capacity and infiltration rates (Meij et al., 2018). Low infiltration rates, high surface runoff, and low root density are considered traits of poor soil physical health as they favour soil erosion and impede plant growth (Dexter et al., 2004).

Soil texture is a crucial factor that affects the balance between water and air in the soil and is generally unaffected by soil management. Consequently, total porosity and bulk density are better indicators for assessing the effects of soil use and management (Yolvubal et al., 2004). Total porosity can be influenced by structural factors (macro/microstructure) and is indirectly related to the soil's texture (proportion of soil particles). Soil management practices can significantly affect structural porosity, particularly macropores, which in turn may alter the soil water retention curve (Dexter, 2004).

Soil is composed of soil particles: sand, silt, and clay. The arrangement of primary soil particles into larger aggregates of varying sizes and shapes, together with the surrounding pore network, determines the soil structure. Dexter (2004) noted that plant growth favours a granular soil structure, as there is a balance between the water and air proportions controlled by micro- and macropores. Soil structure is related to other physical attributes of soil, such as bulk density, porosity, and water infiltration. Moreover, Bini et al. (2013) observed that soil structure is affected by SOM. SOM plays a crucial role in aggregating soil particles together to form stable soil structures. Thus, if soil structure is affected by SOM, then other physical characteristics will also be affected.

Microaggregates are primary aggregated units of soil particles typically between 25-250 µm in diameter, whereas macroaggregates are much larger, and may in fact consist of multiple

microaggregates. Microaggregates are generally more stable and are thus less affected by soil management practices; they are also responsible primarily for stabilising SOC (Six et al., 2004). Soils with high levels of SOM have increased aggregate stability, which also results in increased microbial activity. According to Spaccini et al. (2014), soil carbohydrates account for up to 25% of SOM and act as stabilisers for soil aggregates. Thus, soil aggregates disperse more under more intensive management practices in soils with lower inputs of OM (Qin et al., 2010). Reduced soil aggregates also reduce the number of macropores in the soil. This, in turn, impairs the access of decomposing microbes to SOM (Chodak et al., 2011). Moreover, soil aggregates not only affect nutrient cycling, permeability, and aeration but also harbour microbes. Secretions, cell lysates and mucus produced by organisms, such as earthworms and fungi, act to stick soil aggregates together, further stabilising the soil (Preston et al., 2011). Microbial activity and SOM, therefore, influence the soil structure and, thus, the hydrological processes such as erosion and infiltration rate.

Furthermore, Geisseler et al. (2011) found a correlation between microbial activity and water availability (Geisseler et al., 2011). Thus, reduced microbial activity due to lack of water can lead to degradation of soil functions, such as nutrient cycling and mineralisation of SOM. However, water restrictions affect microbes differently. For instance, fungi prefer drier conditions that allow their hyphae to infiltrate air-filled macropores, whereas bacteria generally prefer moist soil conditions for better movement (Wong et al., 1976; Smith et al., 2017). As a result, soil microbial communities and physical attributes of soil impact each other, and both are affected by SOM (Degen et al., 2000). In addition to physical factors, such as porosity and soil texture, water availability is also dependent on chemical properties, such as SOC.

Microbial activity and diversity play a key role in maintaining essential soil functions, such as C and nutrient cycling. Microbial indicators are more susceptible and responsive to environmental changes like soil management practices compared to physical and chemical indicators (Masto et al., 2009). The soil microbial biomass mineralises the SOM to provide nutrients to the plant through rapid cycling (Sicardi et al., 2004). Furthermore, Kennedy (2013) describes soils with high microbial diversity as more resilient to environmental disturbances due to their functional redundancy.

Soil respiration is a widely used biological indicator of soil health. C sources and nutrients from (S)OM will be mineralised to carbon dioxide (CO_2) and other nutrients through microbial activity (Kennedy et al., 2013). The metabolic quotient is an index used to assess the metabolic activity or efficiency of soil microbes. It provides insights into how efficiently soil microorganisms utilise C for their metabolic processes. The metabolic quotient is often calculated by dividing the rate of CO_2 production by the microbial biomass in the soil, whereby a higher index indicates stressful conditions (Anderson et al., 2007). However, higher index values can also account for easily degradable SOC that stimulates microbial activity. Thus, the quantity and quality of SOM will also determine the mineralisation of C (Babujua et al., 2010). Moreover, soils with high microbial diversity can ensure ecological processes continue after disturbances, making the soil more resilient to change.

Soil enzymes are proteins produced by microorganisms, plants, and other organisms in the soil. These enzymes catalyse biochemical reactions involved in C and nutrient cycling and represent the metabolic level of microbial groups in the soil. Thus, the activity of soil enzymes can provide insights into the functional status and health of the soil ecosystem and can, therefore, be useful indicators of soil health. Soil enzymes can be released by cells into the soil after cell lysis or can be free in the soil as extracellular enzymes (Nayak et al., 2017). Soil enzyme activity is expected to change under different soil management practices (Lines-Kelly et al., 1998). For example, P fertilisers can inhibit phosphatase enzymes in a feedback effect, whereas increased cellulase activities are generally seen in soils with high SOC in the topsoil (Peixoto et al., 2010). Thus, soil microbial communities are affected by soil management, as soil disturbances and additional inputs can change the availability and activity of soil enzymes.

Invertebrates that live in the soil play a key role in establishing relationships with different levels of microbes (Brown et al., 2009). Microorganisms break down and transform SOM into plant-available nutrients; however, this process is made more effective when the SOM is pre-transformed and made more accessible to microbes (Vince et al., 2018). For example, earthworms promote SOM movement across soil horizons, distributing the SOM to different microbial communities (Kostina et al., 2011). The presence or absence of a specific species can be crucial for ecosystem functioning (Schjonning et al., 2004). In cases where

earthworm species have decreased in the soil, plant residues accumulated on the soil surface and were not fully integrated across the soil layers. This happened despite the presence of other functionally similar organisms which were not as proficient in distributing OM as earthworms (Hoogerkamp et al., 1983). Thus, the activity of individual species can limit specific soil processes. As a result, we should consider the species richness index in soil health assessments. Moreover, increased aggregate stability has been seen in soils with higher microbial and earthworm activity (Mader et al., 2002). The movement of these organisms mixes soil particles and produces pores, channels, and other biological compartments that allow the air and water to flow better, which further stimulates microbial activity (Drewry et al., 2008).

Huge variability between different climates and geographical locations can be observed when using biological attributes as indicators for soil health. Therefore, simple indices based on biological parameters do not fully describe soil health or the complexity of soil systems. Taking measurements across different time points and conditions or using statistical methods to account for variability must be considered when conducting soil health tests. Constant monitoring and evaluations of the following physical, chemical, and biological indicators are needed in order to measure soil health reliably. Failure to do so would result in a measurement that does not fully represent the complexities of soil systems. Furthermore, by studying the origins of natural processes, we can better understand the multifaceted connections between the chemical, physical, and biological parameters of the soil. For instance, photosynthesis, which is arguably our most important biological process, allows plants to transform C to carbohydrates for the food web. Plant and animal debris are deposited in the soil, allowing organic C and minerals to be recycled and utilised by other organisms. Biological processes are, therefore, integral for recycling above and below ground C and nutrient mineralisation. Healthy soils have the capacity to maintain these processes sustainably over time.

1.6 Impact of Agricultural Practices on Soil Health.

Farmers employ a range of agricultural practices and inputs to maximise food and fibre production. This includes the use of a range of inputs, thus distorting the natural balance of the soil ecosystem and potentially compromising ecosystem services. As a result,

governments are continuously promoting new legislation and incentives to promote sustainable practices, and one approach is to relate soil functions to natural capital. Sustainable agriculture aims to reduce the detrimental effects of farming by ensuring agricultural production without compromising ecosystem services (Kibblewhite et al., 2007).

Human intervention has altered the natural state of soils, leading to a loss of ecosystem functions. Continuous cultivation and deforestation to create agricultural lands have changed the soil properties that describe soil health. In particular, global SOM losses from land use changes have been increasingly documented over recent years (Beillouin et al., 2021). SOM losses are also correlated with reduced cation exchange capacity, further exacerbating the soil's ability to retain nutrients. Depleted nutrient levels in the soils have considerable effects on the capacity of the soil to deliver ecosystem services. SOM must, therefore, be maintained or preferably enhanced through organic amendment where necessary to restore soil health and thus restore the soil's ability to deliver ecosystem services.

Common agricultural practices artificially mimic natural processes. However, the scale of these actions can significantly affect soil health. For example, the application of pesticides replaces natural biological pest and pathogen control, whilst inorganic fertilisers substitute nutrient cycling. Due to the integrated properties of soils, agricultural practices that aim to replace or modify biological functions will also have indirect effects on other soil functions.

Tillage disturbs the soil with a range of tools to create a suitable seedbed for crops. The process allows crop residues and OM on the soil surface to be incorporated, thereby contributing to increased levels of nutrient cycling. However, as a result of the mechanical disturbance, a decline in SOC and greater N mineralisation and a decline in macrofauna populations (especially earthworms) are observed under tillage systems. Therefore, losses of SOC can be reversed by reducing tillage. Furthermore, Six et al. (1999) noted improved aggregate formation and greater C retention in no-till systems, although others have reported that tillage practices affect the depth distribution of SOC rather than overall stocks. So, while tillage can negatively impact ecosystem services, the effects can be reversed and, in some cases, can even improve crop production. Overall, agricultural intensification

through tillage has facilitated economic savings by increasing efficiency and reducing human labour. However, in recent years, we have begun to understand the long-term effects of tillage on soil ecosystem functioning. As a result, there has been a considerable return to reduced or even no-till practices.

The same can be said when considering the impact of pesticides and fertilisers on soil functioning, particularly when used in excessive quantities. According to Beauchamp (1991), high fertiliser inputs have been correlated with reduced OM quality, and high ammonium concentrations inhibit N fixation and stimulate nitrification, which in turn affects soil microorganisms. Excess N fertiliser application can also contaminate aquatic ecosystems and drinking water and lead to increased nitrous oxide production. A combination of these effects leads to considerable changes in global N cycling. Nitrous oxide has a global warming potential of about 300 times that of carbon dioxide.

Inorganic fertilisers can increase SOM degradation rates by enhancing microbial decomposers that were previously nutrient-limited. As a result, we may see greater SOM decomposition rates as inorganic fertiliser inputs increase (Li et al., 2017). However, many studies have noted an inverse effect. Riggs et al. (2015) observed an inhibitory effect on SOM degradation when inorganic N was added to low N soils. What is more, inorganic fertiliser usually increases the SOM input into soils because it increases plant production and hence results in greater returns of residues to the soil. These conflicting results show the complexity of inorganic fertiliser inputs on C transformations, soil interactions and, thus, soil health.

The unprecedented degradation of our soil health caused by agriculture has led us to question the longevity of current farming practices. Scientists have, therefore, introduced the term ‘sustainable agriculture’ to safeguard agro-ecosystems whilst ensuring sufficient food production. Previously, agriculture was seen to provide a single service, namely the delivery of food and fibre. However, in light of climate change, society requires that all soil functions must be maintained, and this has been echoed by governments across the globe. As a result, current perceptions of agriculture encompass more than just food production but include associated ecosystem services such as water quality and disease control.

Finally, we have to be aware that soil is a finite resource. We must maintain and improve our soils, which will, in turn, increase crop productivity and preserve agricultural systems, which are crucial goals of sustainable agriculture. Present strategies to achieve this include increasing SOM and reducing soil erosion by improving crop diversity and reducing tillage where possible. As our need to increase food production to feed an ever-growing population continues, our agricultural soils must be used intensively but sustainably. The global need for not only improving and restoring soil health but also having metrics to monitor soil health will undoubtedly enhance our understanding of factors that contribute towards sustainable agriculture. By improving the soil's chemical, physical and biological properties and by repairing historical damage, we can ensure agricultural soils can still provide for us in the long term.

1.7 Aims and Objectives

1.7.1 Aim

The main aim of this thesis was to develop a soil health index that was able to quantify soil health for different agricultural land uses and soil textures with the aid of Rothamsted Research's long-term experiments of known history and land management, thereby providing a robust framework for assessing and monitoring soil health in agricultural ecosystems. To achieve this, three main objectives were outlined.

1.7.2 Objective 1

Assess new innovative ways to measure soil health indicators, in particular aggregate stability, and review existing lab methods and evaluate their requirements and functions. Objective 1 is addressed in Chapter 2, where we evaluate the SLAKES application, which is a new method for measuring aggregate stability across different soil types and land management practices.

1.7.3 Objective 2

Investigate the potential of extracellular enzymes as reliable indicators of soil health based on relevance, sensitivity, practicality, and applicability in diverse agricultural contexts. Chapter 3 explores the utility of candidate soil enzymes across various environmental conditions and management practices, with the aim of understanding their role in assessing soil biological activity and functionality within agricultural ecosystems.

1.7.4 Objective 3

Investigate the feasibility of utilising structural equation modelling (SEM), a statistical technique used to test and estimate complex relationships between variables, in the development of a comprehensive soil health index. We aimed to elucidate the complex relationships among soil health indicators and environmental factors, which were discussed in Chapter 4. Chapter 4 also aimed to quantify soil health using SEM to help inform land managers and policymakers about the health of their soils.

1.8 Thesis Structure

Chapter 1: A literature review of the history and concepts of soil health and ways of assessing soil health, as well as discussing the impact of agricultural practices on soil health.

Chapter 2: An experimental study evaluating the use of the SLAKES smartphone application to measure aggregate stability. The chapter compares results from the SLAKES application with the established Le Bissonnais method to understand the sensitivity of the SLAKES application. Here, we aimed to assess the reliability of SLAKES as a method for measuring aggregate stability, to evaluate SLAKES' sensitivity to different land uses and soil types, and to compare SLAKES with the traditional Le Bissonnais method in terms of effectiveness and accuracy.

Chapter 3: An experimental study looking at the use of extracellular enzymes, N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU), as a promising soil health indicator. Also, we provide a validation of which of these soil enzymes, if any, could be used as a comprehensive biological indicator for soil health by examining the relationships between microbial enzyme activity in a range of soils with contrasting chemical and physical

properties. Here, we aimed to understand the importance of extracellular enzymes in monitoring soil health and shortlist and justify the use of particular candidate enzymes (GLU, NAG, and PHO) by comparing enzyme activity within different agricultural systems.

Chapter 4: An experimental study to develop an integrative soil health index by taking measurements of soil properties relevant to soil health, including the physical, chemical, and biological indicators of the soil under contrasting agricultural land uses and soil types. Chapter 4 outlines a framework to develop a soil health index using structural equation modelling incorporating key indicators of soil health and considers the dynamic interactions within agricultural ecosystems, with the ultimate goal of providing actionable insights for optimising soil management practices.

Chapter 5: A summary and discussion of the key findings from both the experimental and analytical chapters of this thesis are provided, reflecting on the current research, discussing its broader implications, and offering recommendations for future research endeavours.

2 Validation of the SLAKES Smartphone Application Against the Le Bissonnais Method Using Long-Term Experiments.

2.1 Abstract

Aggregate stability describes the ability of soil aggregates to remain stable against external forces such as rapid wetting and raindrop impact. Stable aggregates are an indicator of good soil structure, which in turn improves water-holding capacity and protects soil organic matter. Aggregate stability is, therefore, an important physical indicator of soil health. Current methods to measure aggregate stability often involve disrupting soil aggregates in distilled water. These tests are time-consuming, require specialised equipment and are usually done in laboratories.

The Soil Aggregate Stability (SLAKES) smartphone application, developed by the University of Sydney, Australia, quantifies aggregate stability by measuring how quickly soil aggregates disintegrate once submerged in water. The SLAKES application requires three soil aggregates between 2-15 mm in diameter to be placed in a petri dish. Water is added, and the SLAKES app provides a measurement of aggregate stability within 10 minutes using the camera.

To determine the sensitivity of the SLAKES app, we compared its aggregate stability measurements with that of the established “Le Bissonnais” method. Soil samples were taken from fields under fallow, continuous arable cropping, and permanent grass management at two experimental sites with different soil types and different stabilities. The SLAKES app's results were similar to those achieved with the standard Le Bissonnais method. The SLAKES app could differentiate between different management types on clayey soil but was less sensitive when tested with sandy soil. Despite this, we conclude that the SLAKES app is a legitimate, reliable method to measure aggregate stability. The app offers a simple, fast, and cheap alternative to standard laboratory methods, allowing land managers and non-scientists to test the aggregate stability of their soils for themselves.

2.2 Introduction

Aggregate stability is a major physical soil health indicator related to the stability of soil structure – the system of soil solids arranged into aggregates with internal and surrounding pore space (Amézketa 1999). It describes the ability of soil aggregates to remain stable against external forces such as rapid wetting and raindrop impact. While soil structure refers to the heterogenous arrangement of soil in space at a given time, aggregate stability describes the stability of the soil under physical stress and across time (Abivena et al., 2009). Soils with good soil structure can store more water and air, cycle nutrients more efficiently, and reduce the risk of soil erosion. A recent study by Neal et al. (2020) showed that an increase in connected pores supports favourable microbial communities and functions. These microbes process plant and animal residues before incorporating them into the soil organic matter (SOM) pool, which in turn enhances soil structure. The maintenance of good soil structure is, therefore, imperative for sustainable agriculture and the provision of ecosystem services.

Despite aggregate stability being a physical soil property, chemical and biological influences, such as exchangeable cations, microorganisms, and earthworms, play an important role in forming soil aggregates. SOM is also directly related to aggregate stability, as soils are more stable if roots and fungal hyphae living on SOM hold aggregates together. Therefore, soil management practices that increase SOM and support soil organisms can significantly impact aggregate stability. Furthermore, crops and crop rotations that cover the soil surface as much as possible throughout the year protect the soil structure, particularly in surface horizons (FAO, 2003). As a result, land management practices can greatly impact aggregate stability and, therefore, soil health.

Quantifying aggregate stability can, therefore, provide insight into soil structure and inform soil management decisions for land managers and policymakers. Current conventional methods to measure aggregate stability by wet sieving often involve measuring the aggregate size distribution and stability after disrupting soil aggregates in water. Many alternative methodologies have been developed for this purpose, making the comparison of results between different procedures difficult. Although there is no universal method to

measure aggregate stability, the “Le Bissonnais” method (Le Bissonnais, 1996) is as close to a unifying standard as possible as it is accepted by the International Organization for Standardization and its results can be included in soil erosion models (Borrellia et al., 2021; ISO, 2012). The full method combines three treatments of differing wetting conditions and energies: fast wetting, slow wetting, and stirring after pre-wetting, representing the different mechanisms of aggregate disruption, followed by measuring the resulting aggregate size distribution after each treatment. Previous work at Rothamsted Research has proven the fast-wetting treatment to be the most sensitive to separate different soils and different managements (Redmile-Gordon et al., 2020), which has therefore been chosen for this study.

Although the Le Bissonnais method is a standard method producing reliable results, it has its disadvantages. It is very labour-intensive, requiring at least three overnight drying sessions in an oven, uses highly flammable chemicals and a stack of 2-mm to 50-µm aperture sieves. This makes this lab-based technique expensive and time-consuming. It also requires specialised equipment and training, making it an unreasonable method for land managers to use themselves. As a result, a range of studies have tested alternative methods, including the recent development of a phone application (app) called SLAKES (Fajardo & McBratney, 2019; Fajardo et al., 2016). SLAKES is a free smartphone app designed by the University of Sydney that quantifies aggregate stability by measuring how quickly soil aggregates disintegrate once submerged in water. The test itself uses a smartphone, phone stand, Petri dish and tap water and requires very limited training as the application offers simple step-by-step instructions and a video tutorial.

Therefore, the SLAKES app promises to be a viable method for land managers to measure the aggregate stability of their soils themselves. By increasing the availability of quantitative soil structure measurements, this technology could contribute to a better understanding of our global soil resources. But to fulfil this promise, it is essential that the new measurement technique is accurate and reproducible and that the application of the method can be standardised.

Flynn et al. (2020) evaluated the SLAKES app's effectiveness in detecting variations between conventional tillage, no-tillage, and perennial grass management practices, comparing its performance to the Cornell wet aggregate stability test (CWAST). Their results indicated that SLAKES was more sensitive in distinguishing tillage practices than CWAST. Similarly, Rieke et al. (2022) compared the SLAKES app with three other aggregate stability tests. They found that SLAKES had weaker correlations with these tests and limited ability to detect differences in aggregate stability across treatments. Although several studies have compared and evaluated the SLAKES app, its methodology has not been comprehensively tested and compared to other standard aggregate stability measurements.

The aim of this study was to determine whether SLAKES was a valid method for quantifying aggregate stability by comparing its measurements with the established Le Bissonnais laboratory method. We also wanted to assess the limitations of the SLAKES app by using smartphones with differing camera pixel resolutions. We hypothesised that a newer phone with better camera pixel resolution and greater phone memory would perform more consistently during tests and be able to differentiate between field treatments. Furthermore, we also aimed to investigate whether variations in aggregate shape had an impact on the outcomes generated by the SLAKES app. We predicted that results from irregularly shaped aggregates would show significant differences from spherical, regular-shaped aggregates.

2.3 Methods

2.3.1 Long-Term Experiments

We focussed on two long-term field experiments: Highfield Ley-Arable, located at the Rothamsted Farm, Harpenden, Hertfordshire, and Woburn Ley-Arable, located at Woburn Experimental Farm at Woburn, Bedfordshire, both in the southeast of England. The soil at Rothamsted is a silty clay loam developed on London clay-with-flint over chalk, and the soil at Woburn is a sandy loam developed on Lower Greensand. Basic site and soil properties are given in Table 1.

Table 1: Basic site and soil details.

Site	**Highfield Ley-Arable**	**Woburn Ley-Arable**
Location	Rothamsted, Hertfordshire, UK	Woburn, Bedfordshire, UK
Global position (latitude; longitude) (°)	51.8028; –0.3660	51.9991; –0.6167
Altitude (GB Ordnance Datum) (m)	130	99
Maximum annual temperature (°C) [a]	14.0	14.5
Minimum annual temperature (°C) [a]	6.4	6.1
Average annual temperature (°C) [a]	10.2	10.3
Mean annual rainfall (mm) [a]	764	671
Soil		
Soil type (SSEW Soil subgroup) [b]	stagnogleyic paleo-argillic brown earth	typical brown sand
Soil series (SSEW) [c]	Batcombe	Cottenham
Soil type (WRB Reference Soil Group) [d]	Chromic Luvisol	Cambic Arenosol
Sand (g kg^{-1})	150 [e]	708 [f]
Silt (g kg^{-1})	590 [e]	176 [f]
Clay (g kg^{-1})	260 [e]	116 [f]
Soil texture [b]	silty clay loam	sandy loam

[a] 1991–2020 average (Rothamsted Research, 2023)
[b] Soil Survey of England & Wales (Avery, 1980)
[c] Soil Survey of England & Wales (Clayden & Hollis, 1984)
[d] World Reference Base for Soil Resources (IUSS Working Group WRB, 2022)
[e] Avery & Catt, 1995
[f] Catt *et al.*, 1980

Highfield Ley-Arable Experiment

The Highfield Ley-Arable experiment started in 1948 in a field which had been under permanent pasture since 1838 (Lawes & Gilbert, 1885). Some plots stayed under permanent pasture; others went into continuous arable or ley-arable rotation cropping (Jenkinson, 1991) in a randomised complete block design with four replicate blocks. An adjacent long-term bare-fallow area was established in 1959 (Highfield Bare Fallow and the Geescroft Soil Mine), where all vegetation is constantly removed by ploughing 4-6 times a year (Appendix A). We collected soil samples from four blocks (8 plots) from the continuous long-term grass (dominated by perennial rye grass (*Lolium perenne* L.) and clover (*Trifolium* spp.)), and permanent arable (winter wheat (*Triticum aestivum* L.)) treatments in the main Ley-Arable

experiment, and we added to this with 5 plots established on the adjacent bare fallow treatments for this research. Hereafter we refer to these treatments as 'grass', 'arable' and 'fallow', respectively.

Woburn Ley-Arable Experiment

The Woburn Ley-Arable experiment began in 1938 (Johnston, 1972), and has had a varied history (Johnston et al., 2022). The current version of the experiment has, technically, eight 5-year crop rotation treatments, comprising of three years in 'treatment' crops followed by two years under 'test' crops. Four of the treatments have an arable rotation treatment phase comprising winter rye (*Secale cereale* L.) – winter barley (*Hordeum vulgare* L.) - winter oats (*Avena sativa* L.) in years 1, 2 and 3, respectively. The other four treatments are grass leys, either perennial rye grass (*Lolium perenne* L.) with N fertiliser (two treatments) or perennial rye grass – white clover (*Trifolium repens* L.) with no N fertiliser (two treatments). The particular treatments differ, technically, because of their past history (Johnston et al., 2022). Following either the 3-year arable rotation or the 3 years under grass ley, all treatments are then under two years of test crops - winter wheat and winter rye in years 4 and 5, respectively. The treatments are established in paired plots within a block (16 plots for the 8 treatments). There are then 5 blocks representing each year of the 5-year crop rotations in any particular year (80 plots in total) (Appendix A).

We collected soil samples from two blocks (32 plots), one in year 3 of the treatment phase (i.e., plots under winter oats or grass leys), and the other one in year 5 under an arable test crop (winter rye). For the purposes of this study, we grouped all eight arable rotation treatments together, and all eight grass ley treatments together.

Soil Sampling

Soil samples were collected from the top 23 cm of the profile using a spade, transferred in a plastic bag and then transported to the lab on the same day. Soil samples were then stored in a dark, cold room at 5 °C until analysis. Three samples were taken from each plot and treated as internal replicates and analysed independently. The Highfield Ley-Arable experiment involved a total of 39 measurements distributed across the 13 plots outlined

above (5 plots under fallow, 4 plots under arable, and 4 plots under grass), with each plot having three replicates. A further 96 measurements, obtained from 32 plots (8 plots each under arable rotation treatment (year 3), grass leys treatment (year 3), test crop following grass leys treatment (year 5) and test crop following arable rotation treatment (year 5)), with 3 replicates were recorded for the Woburn Ley-Arable experiment.

2.3.2 SLAKES Aggregate Stability Test

The "SLAKES: Soil Aggregate Stability" smartphone app (Fajardo & McBratney, 2019) is available to both Android and Apple users and can be downloaded from the Google Play Store and App Store, respectively. The SLAKES app can analyse the stability of soil aggregates by using image recognition software. First, a reference image of selected soil aggregates is taken. Then, water is added and the app started. The app will measure the area the soil aggregates take up over time and then calculate the aggregate stability.

Soil samples collected from field sites were hand-crumbled to release aggregates 3 to 5 mm in diameter, which were then air-dried for 48 hours. Three spherical aggregates were selected to introduce some uniformity amongst all the aggregate shapes that could be selected. The app recommends selecting spherical soil aggregates as changes in orientation after capturing a reference image may affect results.

Phones were suspended on a ring stand at a height that allowed the camera to capture the whole Petri dish (about 20 cm above the petri dish). Lighting included the laboratory's overhead fluorescent lights and natural light from an adjacent window. The Petri dish was also placed on a white background. Good light and high contrast between the soil aggregates and the background are needed when using the app; otherwise, the SLAKES app mistakes shadows and the background as the soil aggregate. There was also no lighting directly over the Petri dish, as glares in the water could register as a soil aggregate.

For each measurement, three soil aggregates from the same sample were placed in an empty Petri dish. Following the SLAKES app instructions, we then captured a reference image. The aggregates were removed, and the Petri dish was filled with tap water. Tap water

was used to replicate the conditions of the natural environment and is readily available to land managers. The soil aggregates were then transferred back into the water-filled Petri dish, with careful consideration to place the aggregates in a similar position and orientation as the reference image. We pressed the start button, and after 10 minutes, a Slaking Index (SI) value was displayed on the screen. The dispersion of soil aggregates (expressed as the SI) is represented by a Gompertz function that has been fitted to the dissolution data (Equation 1).

Equation 1:

$$SI = ae^{-be^{[1-c.log(t)]}}$$

The Gompertz function is defined by three coefficients: *a, b*, and *c*, where coefficient *a* denotes the maximum predicted dispersion of a soil aggregate, *b* characterises the initial slaking, and *c* represents the ongoing rate of change (Fajardo et al., 2016; Flynn et al., 2020). The SI value observed after the 10-minute imaging period (*t*) is an average of coefficient *a* for all three soil aggregates in the fitted Gompertz function (Flynn et al., 2020).

SLAKES recommends the use of spherical aggregates. However, this may not represent the diversity of soil aggregates present. To test if irregularly shaped (non-spherical) aggregates influenced results, we selected three irregularly shaped aggregates and three regular (spherical) shaped aggregates from each plot for a comparative analysis. We defined regularly shaped aggregates as roughly spherical and relatively round in shape, and irregularly shaped aggregates (rectangular, longer than wide), varying in size and shape.

To test if camera pixel resolution influenced results, four smartphones with differing cameras were used in our study (Table 2). SLAKES was downloaded onto Phone 1 and Phone 3 with a 12-megapixel and 5-megapixel rear camera, respectively. We also used older phone models to test whether the app was compatible and worked effectively on various smartphones (Table 2). All smartphones were reset to their original factory settings before downloading the SLAKES app. Each phone measured soil aggregates collected from each plot selected from the Highfield Ley-Arable and Woburn Ley-Arable Experiments. We also examined the

frequency of Not Applicable (NA) and outlier results (SI above 20) for each measurement of aggregate shapes from Phone 1, with samples taken from the Highfield Ley-Arable experiment only. Furthermore, that SLAKES data collected from Phone 1 were used when comparing results from the SLAKES app with the Le Bissonnais method.

Table 2: List of smartphones used in this study.

Phone model	**Code**	**Phone type**	**Year of purchase**	**Camera Mega-Pixel**	**Internal phone memory (Giga-Bites)**
iPhone 11	Phone 1	iOS	2019	12	128
iPhone 7	Phone 2	iOS	2017	12	32
J3	Phone 3	Android	2015	5	8
Galaxy S6	Phone 4	Android	2017	16	64

We repeated the test for each soil sample three times, so that nine aggregates were measured from each plot.

2.3.3 Le Bissonnais Aggregate Stability Test

We used the fast-wetting component of the Le Bissonnais method to represent the mechanisms of aggregate disruption by slaking, and then measured the resulting aggregate size distribution (Le Bissonnais, 1996).

Soil samples collected from the field were hand-crumbled along existing pores and cracks over a single sieve stack with apertures of 5 mm and 3 mm. Soil aggregates collected in the lower 3 mm sieve (3-5 mm diameter) were then carefully transferred to a drying tin, removing any stones, and placed in the oven at 40 °C overnight. A subsample of 5 g of oven-dried aggregates was weighed and then transferred to a glass beaker filled with 50 mL of deionised water, leaving them to immerse. After 10 minutes, we carefully poured the water, trying to avoid losing any soil or further disruption. This step is not part of the standard Le

Bissonnais method but follows the procedure used by colleagues at Rothamsted Research. Using a wash bottle containing methylated spirits (also known as denatured alcohol), we gently transferred the wet aggregates from the beaker to a 50 µm sieve submerged in a bowl of methylated spirits. The sieve was gently twisted ten times. The sieve containing > 50-µm stable aggregates was then removed from the methylated spirit bowl and left to air dry in the fume cabinet for 2 hours. Once air-dried, we brushed the > 50 µm aggregates onto a drying tin, before placing the drying tin in the oven at 40 °C overnight.

The following day, we made a sieve stack comprising the following sieves from top to bottom: 2-mm, 1-mm, 500-µm, 200-µm, 100-µm, 50-µm, and receiver. We removed the dried > 50 µm aggregates from the oven, transferred the sample to the sieve stack and put a lid on. The sieve stack was gently shaken in a rotary motion for 30 seconds. We then weighed the aggregates remaining in each sieve. Then all > 50 µm aggregate fractions were placed into separate bottles containing sodium hexametaphosphate dispersing solution (2 g L^{-1}) to separate out the sand-sized (> 50-µm) primary particles (Kemper & Rosenau, 1986). After 4 hours of shaking, we assumed that all soil particles were fully dispersed. We thoroughly rinsed the remaining particles over a 50-µm sieve in a sink and transferred the > 50-µm sand-sized particles to a drying tin, which was placed in the oven at 40 °C overnight.

After removing the dried > 50 µm sand particles from the oven, we tipped the sand particles into the same sieve stack described above and put a lid on before gently shaking for 30 seconds. We then weighed the sand particles retained on each sieve. All soil weights were determined to 2 decimal places (0.00 g). The stable aggregate size distribution was then calculated using Equation 2, which can be described by the conventional mean weight diameter (MWD) calculation, having accounted for sand-sized primary particles in each size class:

Equation 2:

$$MWD = \frac{\sum_{i=1}^{n}\left[\overline{\phi}_i.\left(a_{\phi_{i(a+p)}} - p_{\phi_i}\right)\right]}{a_d - \sum_{i=1}^{n} p_{\phi_i}}$$

where: MWD = mean weight diameter (mm)

$\overline{\phi_i}$ = mean aggregate diameter in size class *i* (mm)

$a_{\phi_{i(a+p)}}$ = oven-dry weight of soil (aggregates and sand particles) in size class *i* (g)

p_{ϕ_i} = oven-dry weight of sand particles in size class *i* (g)

a_d = total oven-dry weight of soil (aggregates and sand particles) (g)

2.3.4 Statistical Analysis

Slaking Index (SI) and aggregate size distribution data were analysed by fitting Linear Mixed Models (LLM). Block was assigned as a random factor, treatment as the explanatory variable, and SI and MWD as a response. Analytical pseudo-replicates were treated as random effects to account for the correlation between measurements from the same sample. We then extracted a corresponding analysis of variance (ANOVA) table. Datasets were log-transformed if the residuals were not normally distributed. If the probability (p) of the ANOVA variance ratio statistic (F) was < 0.05, we could conclude that treatment would have a significant effect over the response (SI or MWD). A post hoc test (Tukey HSD) was used to explore where these differences lie. Similarly, ANOVA was performed to see if there were significant differences between phone types and aggregate shape on SI scores. Residuals were plotted against the fitted model values to check for an unequal variance between groups, and a normal quantile–quantile plot was used to check for the normality of the residuals. Linear regression was also used to compare SI with aggregate size distribution. All statistical analysis was conducted in R version 4.3.2 (R Core Team, 2023).

2.4 Results

2.4.1 SLAKES vs. Le Bissonnais

Analysis of MWD on Highfield Ley-Arable soils shows that the Le Bissonnais method was able to detect differences between all treatments. Soil aggregates under grass treatments were significantly more stable than those under arable and fallow treatments ($p < 0.001$), whilst there was only a small but statistically significant difference in soil stability between arable and fallow treatments ($p = 0.006$) (Figure 2A). Whilst the SLAKES app successfully separated the grass treatment with highly stable aggregates from the arable and fallow treatments with less stable aggregates, it was not able to further separate, statistically, the arable and fallow treatments in the way that the Le Bissonnais method could (Figure 2B). The values of the Le Bissonnais method and the SLAKES method are inverse because of the method of calculation. A greater MWD (Le Bissonnais method) and a smaller SI (SLAKES) both indicate soil aggregates of greater stability.

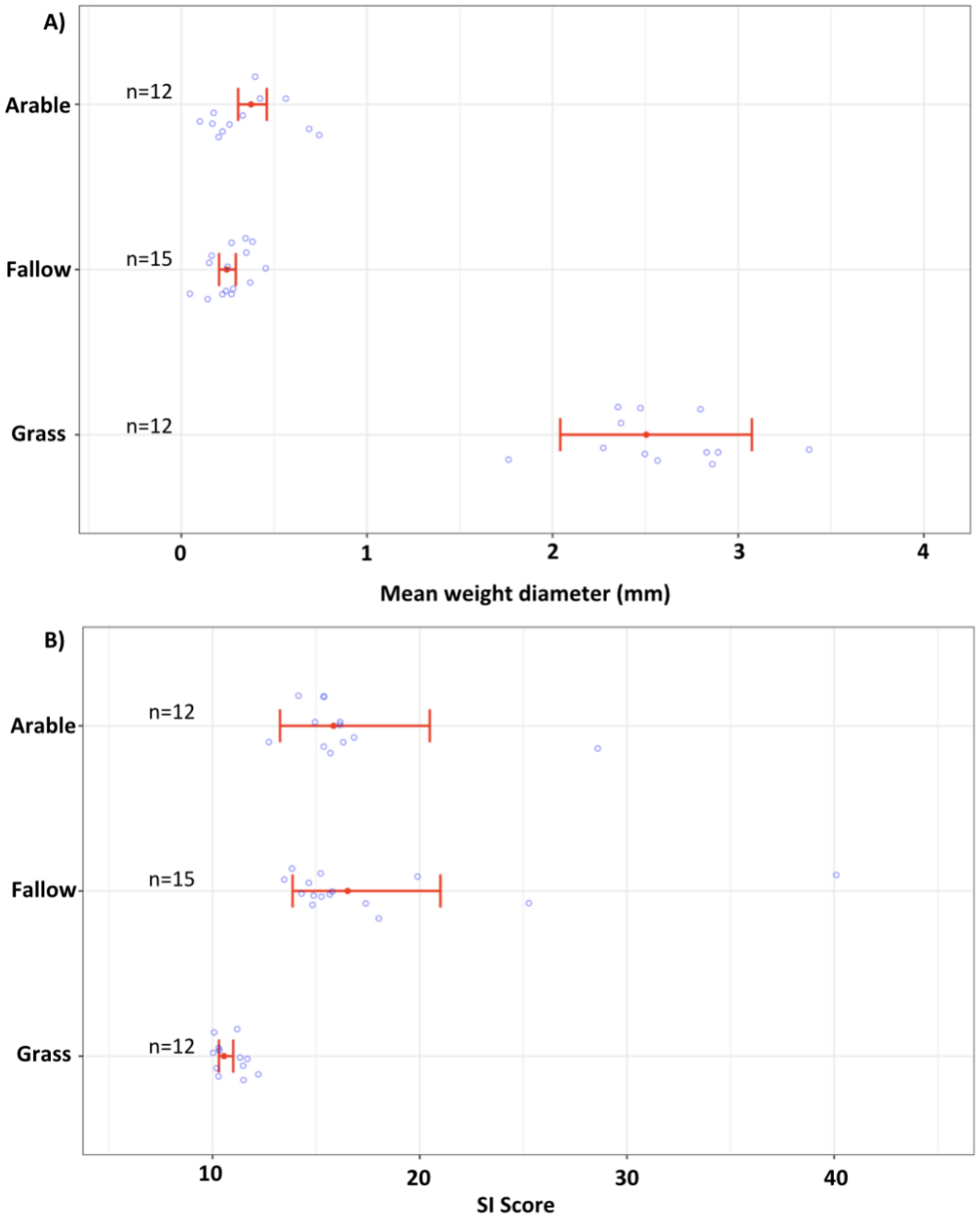


Figure 2: Back transformed treatment means with 95% confidence intervals bar of aggregate stability indices for the A) Le Bissonnais method (mean weight diameter), and B) SLAKES application (SI score) under Grass, Arable and Fallow treatments from the Highfield Ley-Arable long-term experiment.

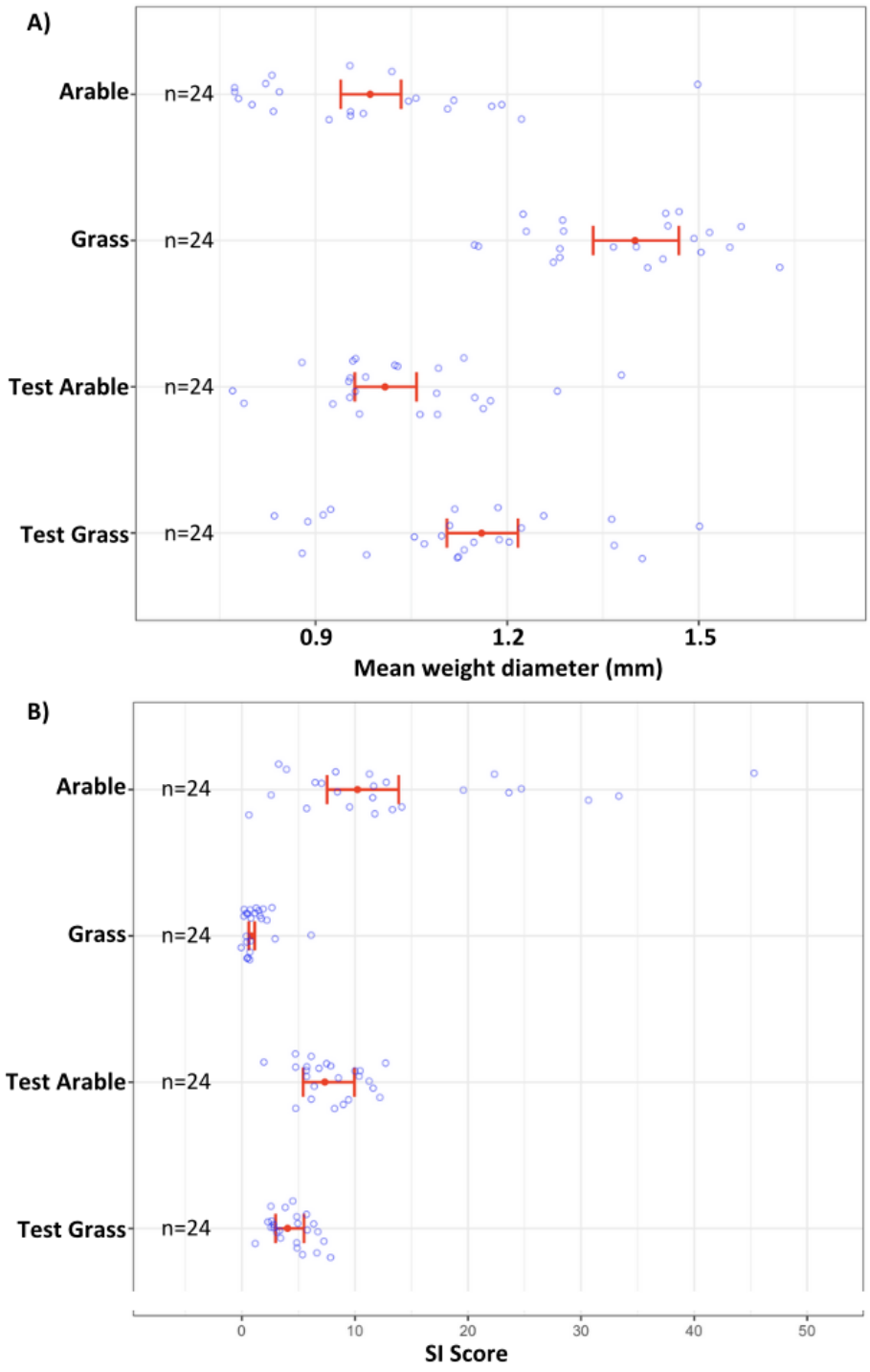


Figure 3: Back transformed treatment means with 95% confidence intervals bar of aggregate stability indices for the A) Le Bissonnais method (mean weight diameter), and B) SLAKES application (SI score) under two crop rotation treatments at two phases: Grass (year 3 – grass ley treatment), Arable (year 3 – arable rotation treatment) and Test Arable and Test Grass (year 5 in test crop following 3 years under arable and grass treatments, respectively) treatments from the Woburn Ley-Arable long-term experiment.

Results showed that the SLAKES application could differentiate between only grass and arable treatments in sandy soils and showed greater variation and spread of data (Figure 3) compared to the clayey soil in the Highfield experiment (Figure 2). Another interesting feature of the Woburn Ley-Arable experiment is that whilst the 'Test Grass' shows some loss of stability in the two years under arable (i.e., compared with 'Grass'), it still remains more stable than treatments which are always under arable crops (comparing with 'Arable' and 'Test Arable').

We also performed a correlation test on log-transformed variables to see if results from the Le Bissonnais method and SLAKES were correlated (Figure 4). The correlation was significant (p-value = <0.001, t=64.12), with a negative slope (-0.51).

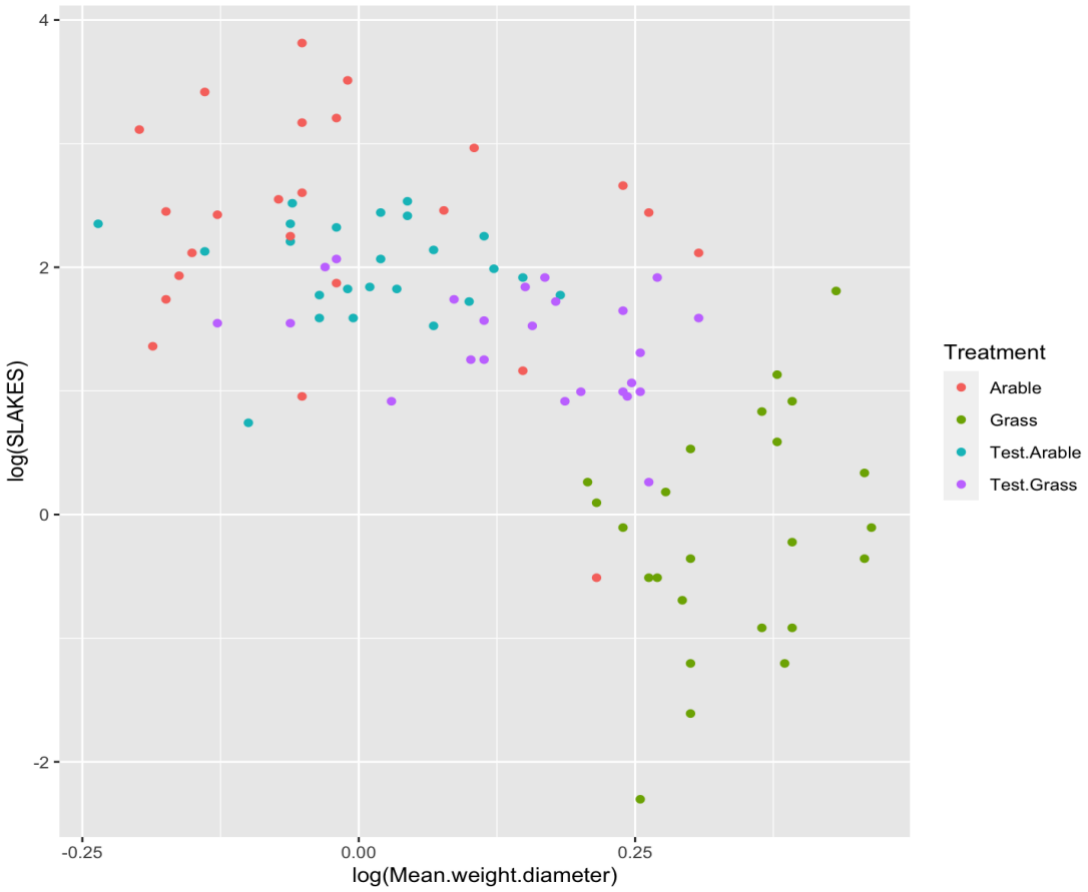


Figure 4: Correlation looking at the SLAKES (Slaking Index score, log-transformed) against the Le Bissonnais method (mean weight diameter, log-transformed) for the Woburn Ley-Arable experiment.

2.4.2 Phone Types and Aggregate Shape

We also ran the SLAKES app on different phones (Phone 1 to 4, see Table 2) with different camera qualities and internal memories and found no significant difference in results between phone types (p value>0.05). However, it is worth noting that Phone 4 could better outline the soil aggregates than Phone 2 (Figure 5). Here, we placed three soil aggregates on a Petri dish under different orientations. Whilst Phone 4 could consistently outline the soil aggregates, Phone 2 was less accurate. We occasionally noted similar issues with Phone 3; however, we observed no such issues with Phone 1. Phone 2's camera lens frequently had issues focusing on the soil aggregates, resulting in a blurry image, which was also the case with Phone 3. Moreover, we found that all phones tended to 'crash' more often after long use. This was more prevalent in older phones such as Phone 2 and Phone 3.

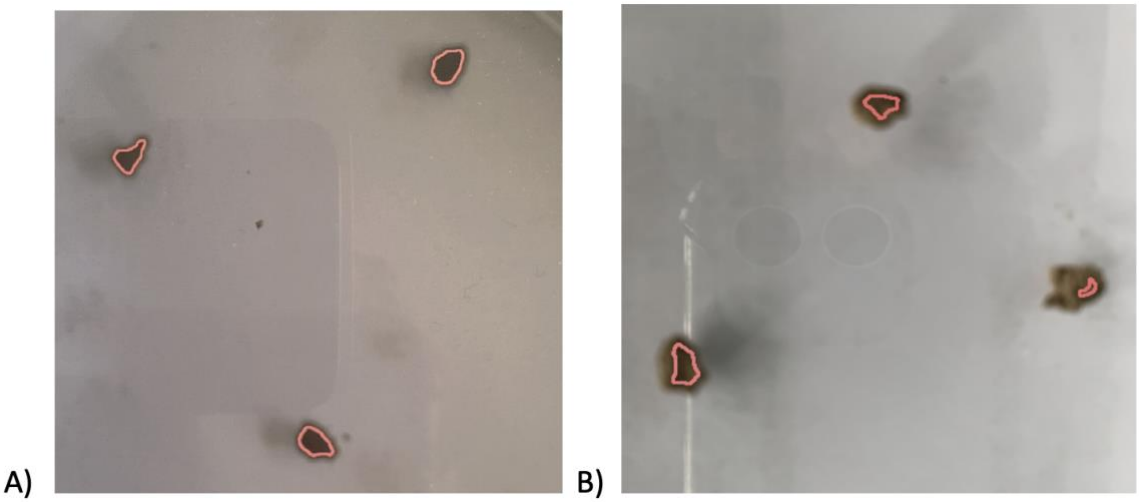


Figure 5: SLAKES reference image of three soil aggregates on a: A) Phone 4 and: B) Phone 2. Each soil aggregate was placed under different orientations. Whilst Phone 4 could accurately outline the soil aggregates, Phone 2 struggled to do so.

We found no significant differences between irregularly shaped aggregates and spherical aggregate shapes. However, we identified irregularly shaped aggregates as a factor that increased the frequency of extreme value, NA results and phone 'crashes' (Table 3).

Table 3: Average frequency of outlier results and phone crashes during SLAKES measurements from the Highfield Ley-Arable experiment for both regularly shaped and irregularly shaped aggregates, taken on Phone 1.

	Regular shaped aggregates	**Irregularly shaped aggregates**
NA	5	11
Phone crashes	2	9

2.5 Discussion

2.5.1 Differences in stability between land use treatments and the ability of SLAKES and Le Bissonnais method to detect them.

Organic matter is closely related to aggregate stability (Tisdall & Oades, 1982), and crops and crop rotations that cover the soil surface can protect soil structure, particularly in surface horizons (Li et al., 2021). Therefore, soil management practices that increase soil organic matter and support soil organisms can greatly impact aggregate stability. As a result, aggregates are more stable under pasture grass treatments and less stable under bare fallow, which has been confirmed by the results of the Le Bissonnais method and the SLAKES app (Dexter et al., 2008). The difference in aggregate stability between arable and permanent fallow are much smaller and could only be detected by the Le Bissonnais method. Similarly, the effect of a previous grass treatment on the aggregate stability of the following test crop treatment could only be detected significantly by the Le Bissonnais method.

Organic matter acts as a binding agent, promoting aggregation and enhancing soil structure stability (Eden et al., 2020). Clay soils generally have greater organic matter content than sandy soils. Thus, sandy soils with smaller organic matter content have fewer binding agents that hold soil particles together, resulting in weaker soil aggregates (Djajadi et al., 2012). This is confirmed by our findings from the Woburn Ley-Arable experiment, where the soil is characterised by a greater sand content and lower stability. Whilst the SLAKES application was able to differentiate between the grass and arable treatments on clayey soils (Figure 2), it was less sensitive on sandy soils compared to the Le Bissonnais method (Figure

3), suggesting that soil type plays a significant role in the SLAKES app functionality. Thus, we suggest that the SLAKES application could be used as an alternative to conventional lab-based wet-sieving techniques when used on fine-textured clayey soils, but operators should be more cautious when using the application on coarse-textured sandier soils due to the losses in sensitivity. But overall, SLAKES was still able to distinguish the effects of most of the different land management treatments in both soil types. Similarly, Adetsu (2021) observed variability in the SLAKES method's sensitivity in detecting changes in soil aggregate stability attributed to soil management practices. Notably, they found that SLAKES demonstrated heightened sensitivity to impacts from manuring and variations in soil types. Our results also followed a similar trend, with soil under grass being the most stable, followed by arable and then fallow treatments (Figure 2).

Results from the SLAKES app are expressed as a slaking index coefficient (SI value), where values <3, 3-7, and >7, indicating high, moderate, and low stability, respectively (Fajardo & McBratney, 2019). All samples taken from grass treatments scored an SI under 3. While samples from arable treatments generally scored under 7, the average SI score was 7.9 for fallow treatments. Our results, therefore, validate this scoring system as our soil samples taken from different land management followed this trend.

Flynn et al. (2020) also tested the sensitivity of the SLAKES application to detect differences between conventional tillage, no-tillage, and perennial grass management practices and compared the results with the Cornell Wet Aggregate Stability Test (CWAST). The CWAST quantifies the proportion of dried aggregates that break apart during a controlled simulated rainfall event, mimicking the energy of vigorous spring rain (Flynn et al., 2020). They found that the SLAKES app quantified the aggregate stability decreasing in the order of perennial grass (strongest), no-tillage, and conventional tillage (weakest) practices. Furthermore, Flynn et al. (2020) noted that SLAKES could differentiate among tillage management practices at a greater sensitivity than the CWAST. In contrast, we found that our lab-based method was more sensitive in detecting differences between land management compared to SLAKES. It is worth highlighting that the mechanisms triggering disaggregation in these tests vary, which could potentially help explain the differences in results (Flynn et al., 2020). Moreover, soil aggregates were subjected to two different degrees of drying. Aggregates intended for the SLAKES app were air-dried, whereas those for

the Le Bissonnais method were oven-dried (but only at 40 °C). This distinction in drying methods may account for the difference in sensitivity observed between the two methods.

Another study by Rieke et al. (2022) compared the SLAKES app with three other aggregate stability tests. They found that SLAKES was correlated the least with other aggregate stability tests and found that the SLAKES app did not effectively find differences in the aggregate stability across different treatments. However, our results found that SLAKES and the Le Bissonnais method were significantly correlated (Figure 4).

2.5.2 Frequency of outlier and what might influence them

We also observed many outlier values from the SLAKES app, which were predominantly seen in unstable soils under the fallow and arable treatments, with values as high as 14,000. We initially did not remove any outliers from our analyses as one aim of this study was to identify whether the SLAKES app could be used by non-scientists, and we argued that land managers might not be able to interpret outlier values correctly. However, we recommend removing values above 20 and repeating the analysis in these cases. Another option to increase the precision of the mean would be to increase the number of measurements. However, in our study, we already conducted 24 individual measurements, which is possibly close to what would be feasible for a field.

The analysis of the frequency of outliers identified irregularly shaped aggregates as a factor that increased the frequency of these extreme results (Table 3). Irregular-shaped aggregates are not symmetrical, so more care is needed when placing the aggregate back into the Petri dish of water at the same angle after taking a reference image. This is important as the surface area outlined could be reduced after the reference image is captured, which can lead to a result with a very high number, an error message (NA) or the application crashing. Moreover, it is important to place the aggregate in the same orientation as the reference image to avoid a misleading SI. This is evident in Figure 5, where the same three soil aggregates were placed in different orientations. Note that soil aggregates should always be removed from the Petri dish before adding water; this is because the soil aggregates may shift out of the frame or move too close together even if water is carefully added. Once this happens, any attempts to move the soil aggregates into their original position will affect the results.

We also found that outlier results could be due to one or more soil aggregates not being recognised when the app is running after taking the initial reference image, which in turn influences the final SI. For example, if the app recognises a shadow as an aggregate whilst the app is running, this can lead to an outlier result. The SI is calculated by averaging all aggregates in each measurement. So, if one aggregate is captured incorrectly, the SI is skewed for all aggregates in the test as a result. Using image recognition R code, users can view the area (pixels) of individual aggregates over time and potentially exclude individual aggregate readings that are inconsistent within each measurement (Fajardo et al., 2016). Nonetheless, doing so would require knowledge and some level of expertise in R studio, which we argue not all land managers will have.

Furthermore, the occurrence of these outlier values and phone 'crashes' (NA) increased with phone use over time. We believe this is due to each SLAKES measurement requiring 5 to 6 GB of physical memory, depending on the aggregate size and shape (Fajardo et al., 2016). Initially, we took concurrent measurements, but after observing an increased frequency of outlier and NA results, phone crashes, and overheated phones, we decided it was best to leave up to 10-minute gaps between each measurement. This could further explain why we observed better SLAKES performance in phones with higher memory and CPUs, such as Phone 1 (Table 3).

Sometimes, it was difficult to focus the camera on the soil aggregates (Figure 5), which seemed to be largely due to inadequate lighting. When using a ring light, the camera lens was able to focus on the soil aggregates better but then had issues with shadows and glares. Obour et al. (2023) conducted a comparison between the SLAKES application and the dry aggregate stability (DAS) method. They also noted instances where the SLAKES method had difficulties in accurately detecting soil aggregates and effectively capturing the disaggregation process during stability measurements.

Therefore, the user should first establish a good equipment set-up to realise the benefits of the SLAKES method as a quick and easy method for evaluating soil aggregate stability. So, whilst the SLAKES app can be operated outside of the lab, proper setup conditions are still required. Good lighting, a stable, flat surface, and a white high-contrast background are needed for best results. Consequently, we do not recommend the SLAKES app as an in-field method, and we highly recommend its use indoors to minimise the effects of interfering factors.

2.6 Conclusions

In summary, the SLAKES app can identify stable and unstable soil aggregates and differentiate between management types. However, it is less sensitive and less precise than the Le Bissonnais method. As expected, SLAKES was able to identify pasture grass as the management producing the most stable aggregates but was not able to differentiate between arable and bare fallow treatments. It was also less able to differentiate between treatments in a sandy soil compared to a clayey soil. However, the method did produce reasonable average values that could allow a simple quantitative evaluation of aggregate stability. Thus, allowing us to take a higher number of measurements from the same sites.

We found that the app was particularly sensitive to light. It can mistake shadows and light glares on the water as aggregates. As a result, if the experiment is performed without adequate lighting and a high-contrast background, the application will fail. We also discovered that there was a lot of variation across each reading. This was largely found when aggregates were not spherical and were irregular in shape. Careful consideration is therefore needed when selecting the soil aggregates. However, this may introduce an additional layer of bias, as we deliberately choose specific aggregates, potentially leading to non-representative measurements.

We believe the SLAKES app is a legitimate and reliable method for measuring aggregate stability. The SLAKES app offers a simple, fast, and cheap alternative to standard conventional laboratory methods, allowing non-scientists to test the quality of their soils themselves. Also, by increasing the availability of quantitative soil structure measurements, this technology could contribute to a better understanding of our soils.

3 Extracellular Enzymes as Promising Soil Health Indicators: Assessing Extracellular Enzyme Response to Different Land Uses, Organic Amendment Additions and Soil Textures Using Long-Term Experiments.

3.1 Abstract

Extracellular enzymes (EE) play a crucial role in soil organic matter (SOM) decomposition and nutrient cycling and are known indicators of soil health. However, it is poorly understood how these enzymes respond to different land uses, and their relationships to other soil properties have not been extensively researched.

Long-term experiments at Rothamsted's Woburn and Harpenden sites in the UK were used to evaluate how different management practices affect enzyme activity involved in carbon (C), nitrogen (N) and phosphorus (P) cycling in the soil. Samples were collected from soils with different organic treatments such as straw, farmyard manure (FYM), compost additions, cover crops and permanent grass cover, and assessed for the activities of three soil enzymes: N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU).

Our objective in this paper was to provide a validation of which soil enzymes, if any, could be used as a comprehensive biological indicator for soil health by examining the relationships between microbial enzyme activity in a range of soils with contrasting chemical and physical properties. Furthermore, we wanted to understand the effects of land use and organic amendments on the enzymatic activity of NAG, GLU and PHO in the soil. Our results show that land uses with the least amount of soil disturbances had greater biological activity. Grass treatments relative to all other plots showed the greatest levels of enzymatic activity, followed by arable and fallow, respectively. Soil EE activity correlated with other observed soil health indicators. In particular, NAG and GLU showed a positive correlation with total C and total N, whereas PHO was correlated with inorganic-P (PO_4-P) levels.

Investigating the interactions of important enzymes with soil characteristics and SOC can help us to better understand the health of our soils. Studies on long-term experiments with known histories and large datasets can help us. SOC tends to decrease during land use

changes from natural ecosystems to agricultural systems. Therefore, it is imperative that agricultural lands find ways to increase and/or maintain SOM in the soil.

3.2 Introduction

We are seeing an unprecedented demand for food and fibre as our population continues to grow. Nearly all potential arable land is already in production (FAO, 2020). Maintaining and improving the productive capacity of our soils is essential for human survival. Therefore, healthy soils remain an essential element of agriculture (FAO, 2020). The topsoil, typically the top 15 to 30 cm, is rich in soil organic matter (SOM), nutrients, and microorganisms, making it the most fertile layer of the soil (Kunlanit et al., 2020). In this layer, microorganisms and fauna are primary agents, driving all dynamic processes in soil ecosystems. Microorganisms closely interact with their surroundings due to their high surface area-to-volume ratio. As a result, they can respond quickly to environmental changes, especially when compared to higher organisms. In some cases, changes in microbial activity can cause chemical and physical changes in the soil. Microorganisms can, therefore, give an integrated measure of soil health and provide early signs of soil degradation or, indeed, improvement (Pankhurst et al., 1995). Furthermore, some chemicals that directly impact soil health depend on microbial activities; as a result, soil enzymes can act as important soil health indicators.

Soil enzymes are vital for many biogeochemical processes that influence the overall health of our soils. Extracellular enzymes (EE) are secreted by microorganisms and plant roots and can be found especially in the rhizosphere, the soil zone intimately surrounding and directly influenced by plant roots. They can also be found in pores between soil aggregates, around decaying SOM and on microbial cell surfaces. EE are vital to nutrient cycling and SOM decomposition by catabolising highly complex and diverse organic compounds into readily available nutrients that facilitate plant growth and microbial activity (Burns., 1979). The different types of EE found in the soil are influenced by the content and composition of SOM, which varies with soil type and land management, and can alter soil metabolic processes (Keller et al., 2023). EE can, therefore, indicate microbial community diversity and metabolic activities and help maintain beneficial physical and chemical properties of the soil,

as well as soil fertility, ecology, and health (Gessner et al., 2010). Given their role in mineralising carbon (C), nitrogen (N) and phosphorus (P), as well as stabilising SOM, EE activities are promising indicators of soil health.

SOM decomposition is an essential process in soil ecosystems with implications for C sequestration and nutrient cycling; EE involved in SOM decomposition is, therefore, of particular interest. Cellulase enzymes catalyse the breakdown of glycosidic bonds in cellulose, the main component of plant cell walls. At least three enzymes are needed to degrade cellulose completely into glucose: endo-β-1,4-glucanase, exo-β-1,4-glucanase, and β-glucosidase (Uzuner et al., 2019). Endoglucanase randomly cleaves internal glycosidic bonds within the crystalline structure of cellulose, producing non-reducing ends of carbohydrate chains. Oligosaccharides, such as cellobiose, are removed from these non-reducing chains by exoglucanase. Finally, β-glucosidase (GLU) hydrolyses these oligosaccharides into glucose (Wu et al., 2018).

GLU is a common soil enzyme that catalyses the hydrolysis and decomposition of plant debris in the soil ecosystem (Martinez et al., 1997). GLU regulates the supply of glucose, a necessary C energy source for the activity and growth of many soil microorganisms, and a rate-limiting step in cellulose degradation (Turner et al., 2002). As a result, GLU is a promising soil health indicator as it can provide information on SOM degradation, biological activity, and the effects of land management on the soil. Furthermore, according to Acosta et al. (2000), GLU is sensitive to pH changes, allowing the enzyme to be used as a biochemical indicator for environmental changes resulting from soil acidification (Ali et al., 2019). A better understanding of GLU and other properties that influence them will undoubtedly contribute towards our understanding of soil health.

Chitin is the main component in fungal cell walls and the exoskeletons of arthropods, providing rigidity, strength, and structural support. After cellulose, chitin is the most abundant polysaccharide in nature and contributes towards the organic N found in the soil (Paul et al., 1996). This N-containing polymer is comprised of linear chains of N-acetylglucosamine units linked together by β-1,4-glycosidic bonds (Chen et al., 2010). The main enzymes involved in the complete breakdown of chitin are endo-chitinase, chitobiase

and N-acetyl-β-glucosaminidase. Endo-chitinase cleaves glycosidic bonds within the chitin polymer to release acetyl-glucosamine. Chitobiase hydrolyses the non-reducing end of chito-oligosaccharides to release chitobiose. N-acetyl-β-glucosaminidase (NAG) hydrolyses adjacent N-acetylglucosamine units from the non-reducing ends in chito-oligosaccharides (Lee et al., 2021; Parham et al., 2000). Therefore, NAG is an essential enzyme released by plants and microorganisms that is of great importance in agricultural systems (Deshpande, 1986). Its presence has demonstrated effectiveness in controlling many soil-borne diseases by degrading the cell walls of pathogenic fungi (Singh et al., 2013). Previous research has identified the application of NAG in combating pests and pathogens to increase crop yields and plant growth and, therefore, maintain soil health. However, the role of NAG as a potential indicator of soil health has been widely overlooked.

In soil ecosystems, phosphatases are a group of enzymes that play critical roles in P-cycling. There are two main types of soil phosphatases. Alkaline and acid phosphatases (PHO) work optimally under alkaline or acidic conditions, respectively. Karthikeyan et al. (2002) state that PHO affects plant growth and development by directly influencing the availability of P during low P conditions. Plant roots increase the secretion of the enzymes when signalling indicates P deficiency in the soil. These enzymes then hydrolyse phosphate ester bonds to release inorganic-P from organic-P compounds, increasing the abundance of P. The low substrate specificity of the enzyme allows it to act on a range of structurally related substrates (Alef et al., 1995). Thus, phosphatase enzymes are correlated to plant growth and P-stress and play an active role in P-cycling.

The activities of soil PHO are influenced by soil pH, microbial activity, and the availability of organic P compounds. Understanding these factors could help farmers actively manage P inputs to contribute towards sustainable soil management practices in agricultural systems. Furthermore, understanding the dynamics of enzyme activities in soil ecosystems is crucial for predicting their interactions with nutrient uptake and plant growth and, therefore, soil health.

This overview indicates that EE are sensitive to changes in the environment (Gomez et al., 2020). Monitoring the EE activities in soils can provide valuable insights into the effects of

land use practices, climate change and agriculture on soil ecosystems. Understanding soil enzyme dynamics is important for sustainable agriculture as it could allow land managers to assess soil health, make informed decisions about nutrient management, and develop strategies for mitigating the impact of environmental changes on soil ecosystems. But how soil enzymes respond to different land uses and organic amendments in agricultural systems is not fully understood. Therefore, our objective in this study was to understand the effects of land use and organic amendments on the enzymatic activity of NAG, GLU and PHO in the soil. We also examined the relationships between microbial enzyme activity in a range of soils with contrasting chemical and physical properties, which would further aid our understanding of soil enzyme performances in relation to soil health. Furthermore, using our findings we then assessed which soil enzymes could be used as candidate biological indicators for soil health.

3.3 Methods

3.3.1 Site Locations and Experimental Design

Highfield Ley-Arable Experiment

The Highfield Ley-Arable long-term experiment located at Rothamsted Research in Harpenden, UK (geolocation: 51.802777, -0.366025), started in 1948 in a field that had been in permanent grass since 1838 (e-RA, 2023). In all, six treatments were established in a randomised complete block design with four replicate blocks. The treatments included permanent grass (i.e., continuation of the former land use), arable rotations, and alternating grass ley-arable rotations. In 1959, an area adjacent to the Highfield Ley-Arable experiment at Rothamsted was ploughed and maintained bare since through cultivation; this is the Highfield Bare Fallow. A similar area of long-term bare fallow (the 'Geescroft Soil Mine') is located adjacent to the Highfield Bare Fallow. We focused on the plots under long-term (since 1948 or 1959) grass, arable, and bare fallow treatments from these three experimental areas for this research. These treatments are all continuous with no rotation ('arable' is continuous winter wheat since 2008). There were four replicate plots each of the grass and arable treatments on the Highfield Ley-Arable experiment, and three and two

replicate plots of fallow on the Highfield Bare Fallow and Geescroft Soil Mine areas, respectively, totalling 13 plots in total.

Table 4: Current treatment components in the Highfield Ley-Arable (HL-A), Highfield Bare Fallow (HBF) and Geescroft Soil Mine (GSM) experiments at Rothamsted Research, Harpenden, which were sampled for this study.

Treatment	**Description of treatment**
Grass	HL-A: Grass/clover ley since before 1948; continuous
Arable	HL-A: Arable since 1948; continuous winter wheat since 2008
Fallow	HBF and GSM: Permanent bare fallow since 1959; continuous

Woburn Organic Manuring Experiment

The Woburn Organic Manuring Experiment started in 1964 on a sandy loam soil at the Woburn Experimental Farm in Bedfordshire, UK (geolocation: 51.999805, -0.616036) to test the effects of different types of organic matter (OM) inputs on SOM and crop yields. The experiment has had three distinct phases of OM input, always with 8 treatments in a randomised complete block design with four replicate blocks (32 plots in all). Initially, six organic treatments (FYM, peat, straw, green manure and two grass leys) were compared with two mineral-fertilizer-only treatments. In 2003, the third and current treatment phase started, with 8 treatments (Table 5). An arable rotation (winter rye, spring barley, winter beans, winter wheat, forage maize) was started on seven treatments; the eighth treatment was sown to a grass/clover ley. The seven treatments under the arable rotation are split into 6 split plots to receive 6 levels of inorganic N inputs (except for the winter beans phase, which does not receive inorganic N), which rotates annually. In a seven-year period (accounting for the year of no N under winter beans), each split plot will have received the same amount of mineral N, and soil sampling conventionally is done across the whole main plot area to capture a representative sample across all split plots.

Table 5: Current treatment components (since 2003) in the Woburn Organic Manuring experiment. Note that there are two control (F) treatments – one is a long-term inorganically-fertilised treatment, and the other has previously received organic amendment in a previous phase of the experiment.

Treatments since 2003	
Code	**Treatment**
Control (two treatments)	Inorganic fertiliser only (no organic amendment)
Straw	Chopped straw at 7.5 t ha^{-1}
Cover crop	Cover crop (white mustard) when under a spring crop
Compost	Compost at 40 t ha^{-1}
FYM10 (DG10)	Farmyard manure (FYM) at 10 t ha^{-1}
FYM25 (DG25)	FYM at 25 t ha^{-1}
Grass	Grass/clover leys

3.3.2 Sample Collection

Soil samples were taken mid-season 2022, 2 months after organic amendment application. An auger was used to collect four topsoil samples (0 to 25 cm) from each plot (13 plots in the Highfield Ley-Arable, Highfield Bare Fallow and Geescroft Soil Mine experiments, and 32 plots in the Woburn Organic Manuring experiment). The auger was sterilised between each plot using 70% ethanol. The four samples from each plot were then homogenised, resulting in one sample per plot (45 samples in total).

For the enzyme assays, a subset of the homogenised soil sample from each plot was placed into a sterile vial (50 mL) and stored at 4 °C overnight before being stored at -80 °C until the day of analysis. Samples were allowed to defrost before conducting the enzyme assay. For the chemical analysis, the remaining homogenised soil sample were air-dried for 2 days, and sieved to 2 mm followed by measuring total C, total N, inorganic-P ($P\text{-}PO_4$) and pH.

3.3.3 Chemical Soil Analysis

Total C and Total N

The LECO TruMac Combustion Analyser (LECO Corporation, St. Joseph, MI, USA) is a fully automatic instrument used for the determination of total N and C in soils and plant material, based on a modified version of the `Dumas' digestion method (Dumas, 1831).

Samples were weighed into ceramic `boats' or tin foil cups and placed on an auto-sampler. The sample then entered the combustion chamber, where the furnace and flow of oxygen gas caused the sample to combust. The combustion process converts any elemental C, S, and N to CO_2, SO_2, N_2, and NOx. These gasses are then passed through two anhydrone tubes to remove H_2O, a particle filter, and collected in a ballast tank. The gas was left to equilibrate before being released into an aliquot loop and through the infrared cells, where carbon was detected. Gas passed from the aliquot loop to the catalyst heater where NOx was reduced to N_2, then through Lecosorb to remove CO_2 and anhydrone to remove H_2O. The remaining N_2 and helium carrier gas flowed through a thermal conductivity cell where the nitrogen was measured.

PO_4-P

Olsen P measurements as an indicator of soil P availability were determined in extractions from 5 g of air-dried, <2-mm soil with 0.5 M sodium bicarbonate at pH = 8.5. Soil samples were shaken for 30 min on an orbital shaker (120 rpm, 20 °C) and filtered through Whatman 42 filter paper. Phosphorus in the bicarbonate solution was determined by a phospho-molybdenum blue method on the Skalar SAN^{PLUS} System (Skalar Analytical B.V., Breda, The Netherlands), a continuous colourimetric flow analysis. Refer to Blitz et al. (1948) for a more detailed methodology.

pH

Soil pH is a measure of the hydrogen ion activity in soil solution (pH = $-\log_{10}[H^+]$). The electrometric pH reading is a product of complex electrode interactions between the electrode and the soil suspension; differences in soil:water extraction ratio, electrolyte

concentration of the soil suspension, and spatial placement of the electrode can all affect this reading (Mclean 1982).

Two subsamples (15 g each) of a sieved (<2 mm) soil sample are placed into replicate centrifuge tubes. Tubes were capped to avoid moisture loss and, as necessary, stored in the refrigerator (5 °C) until ready for analysis within a day. Using a pipet dispenser or graduated cylinder, 30 mL deionised water was added to each tube to achieve a soil: water ratio of 1:2. Tubes were capped and shaken for a few seconds; the cap was then removed, and the slurry was allowed to equilibrate with atmospheric CO_2 and warmed to room temperature for at least 30 minutes. The electrode of a pH meter, standardised at pH 7 and 4, was placed into the solution while gently swirling the slurry in the tube. pH was measured to the nearest 0.01. Between samples, the electrode was rinsed with deionised water.

3.3.4 Enzyme Assays

We used an extract-based fluorimetric microplate assay with methylumbelliferone (MUB) as a fluorescence indicator to measure the activities of β-Glucosidase (GLU), N-acetyl-β-glucosaminidase (NAG) and acid phosphatase (PHO) using the substrates 4-MUB-β-D-Glucoside, 4-MUB-N-acetyl-β-glucosaminidase, and 4-MUB-phosphate, respectively (Criquet et al., 1999).

Table 6: Summary of the roles of soil extracellular enzymes that were used in our study.

Enzyme Type	**Enzyme Reaction**	**Enzyme name**	**Enzyme notation**	**Function of enzyme**
β-glucosidase	Cellobiose hydrolysis	β-glucosidase	GLU	C-cycling
Chitinase	Chitin hydrolysis	N-acetyl-β-glucosaminidase	NAG	C-cycling N-cycling
Phosphatase	Mineralises organic phosphate	Acid phosphatase	PHO	P-cycling

For the EE extraction, 3 g fresh soil was added to 50 mL of extraction solution (22.2 g $CaCl_2$, 20 g polyvinylpolypyrrolidone and 0.5 mL Tween 80 in 1 L dH_2O) and shaken at 150 oscillations min^{-1} for 1.5 hours. The samples were then centrifuged at 10,000 rpm for 10 minutes at 4 °C, and the supernatant was filtered through 1.2 µm filters (Whatmann™ GF/C) and dialysed for 12 hours using cellulose dialysis tubes (10-12 kDa cut-off) coated in polyethylene glycol (PEG). Following dialysis, the enzymes were recovered in 10 mL phosphate buffer (0.378 g Na_2HPO_4, 6.9 g KH_2PO_4 in 800 mL of dH_2O; pH 5.6), and the solution was separated into two 5 mL aliquots. One aliquot (live sample) was stored at 4 °C, and the second aliquot was pasteurised for 3 hours at 100 °C to deactivate enzymes, creating a positive control.

Assays were conducted in 96-well black microplates using 38 µL extract and 250 µL substrate (200 µM dilution). For each sample, eight wells (four analytical replicates each for live and control solutions) were filled, and the plates were incubated for 3 hours at 25 °C. Fluorescence was determined on a microplate reader (Fluostar Omega, BMG Labtech, Ortenberg, Germany) at 355 excitation, 460 emission and 975 gain.

A calibration curve was established from nine concentrations of MUB diluted in dH_2O (0-11.5 µmol L^{-1}, eight wells per concentration), read at the same settings. For each sample, the mean 'control' fluorescence was subtracted from the mean 'live' fluorescence and the resulting value (*F*) was then converted to activity (nmol h^{-1} g^{-1}) following Equation 3.

Equation 3:

$$\text{Enzyme activity} = \frac{(F/V)}{c \times t \times M \times D}$$

Where *F* is the corrected mean fluorescence value for a given sample, *V* is the assay well volume (288 µL), *c* is the R^2 value of the calibration curve, *t* is the incubation time in hours, *M* is the sample mass in g dry weight, and *D* is the dilution coefficient 0.2 (50 mL initial extract volume concentrated to 10 mL final volume).

3.3.5 Statistical Analysis

Enzymatic activities of NAG, PHO and GLU were analysed by Type II ANOVA on the linear mixed-effects model using the Kenward-Roger method for computing degrees of freedom to assess the significance of the fixed effects in a mixed-effects model. Correlation coefficients were also used to compare the EE activities with the chemical properties of the soil. Residuals were plotted against the fitted model values to check for an unequal variance for both the linear model and the ANOVAs, and a normal quantile–quantile plot was used to check for normality of the residuals. EE data was log-transformed. All statistical analysis was conducted on R Studio (Version 4.0.4).

3.4 Results

3.4.1 Land Use Change and EE Activity

Samples taken from the Highfield Ley-Arable long-term experiment were used to understand how sensitive each of our chosen EE activities was to differentiate between each land use (grass, arable and fallow). We expected greater soil health in plots with the least amount of soil disturbances. Uncultivated soils under perennial vegetations and soils that receive organic amendments would show increased levels of enzymatic activity compared with soils that are cultivated in annual crop systems and soils that receive only inorganic amendments. We predict that soil EE activity would correlate with other observed soil health indicators. In particular, NAG and GLU would show a positive correlation with total C and potentially with total N, whereas PHO would positively correlate with inorganic-P (PO_4-P) levels.

Our results show that NAG activity was affected by land use, as there were clear differences between each treatment (Figure 6). The NAG activity showed significant differences between the grass and the other treatments (F value: 12.932, p-value 0.0017) but was not able to statistically differentiate between arable and fallow. Likewise, PHO activity was greatest under grass treatments and lowest in fallow treatments (Figure 7). Whilst clear groupings were observed between each treatment, PHO activity was only significantly greater under grass (F value 9.271, p-value 0.00000317), and there were no significant differences between the arable and fallow treatments. Finally, our results show that GLU

activity was most sensitive to land uses, and we were able to detect significant differences between all three treatments (F value: 83.142, p-value 5.874e-07), with grass having the largest GLU activity (Figure 8). Thus, GLU was the most sensitive to land use, whilst PHO and NAG activities could not significantly distinguish between arable and fallow land use.

Overall, NAG, PHO and GLU showed the greatest activity levels under grass, followed by arable, and then fallow. The results show that NAG, PHO and GLU activities were always significantly different between grass and the other land use treatments. Differences in EE activity between arable and fallow land uses showed a consistent trend (arable > fallow) but were not always significantly different.

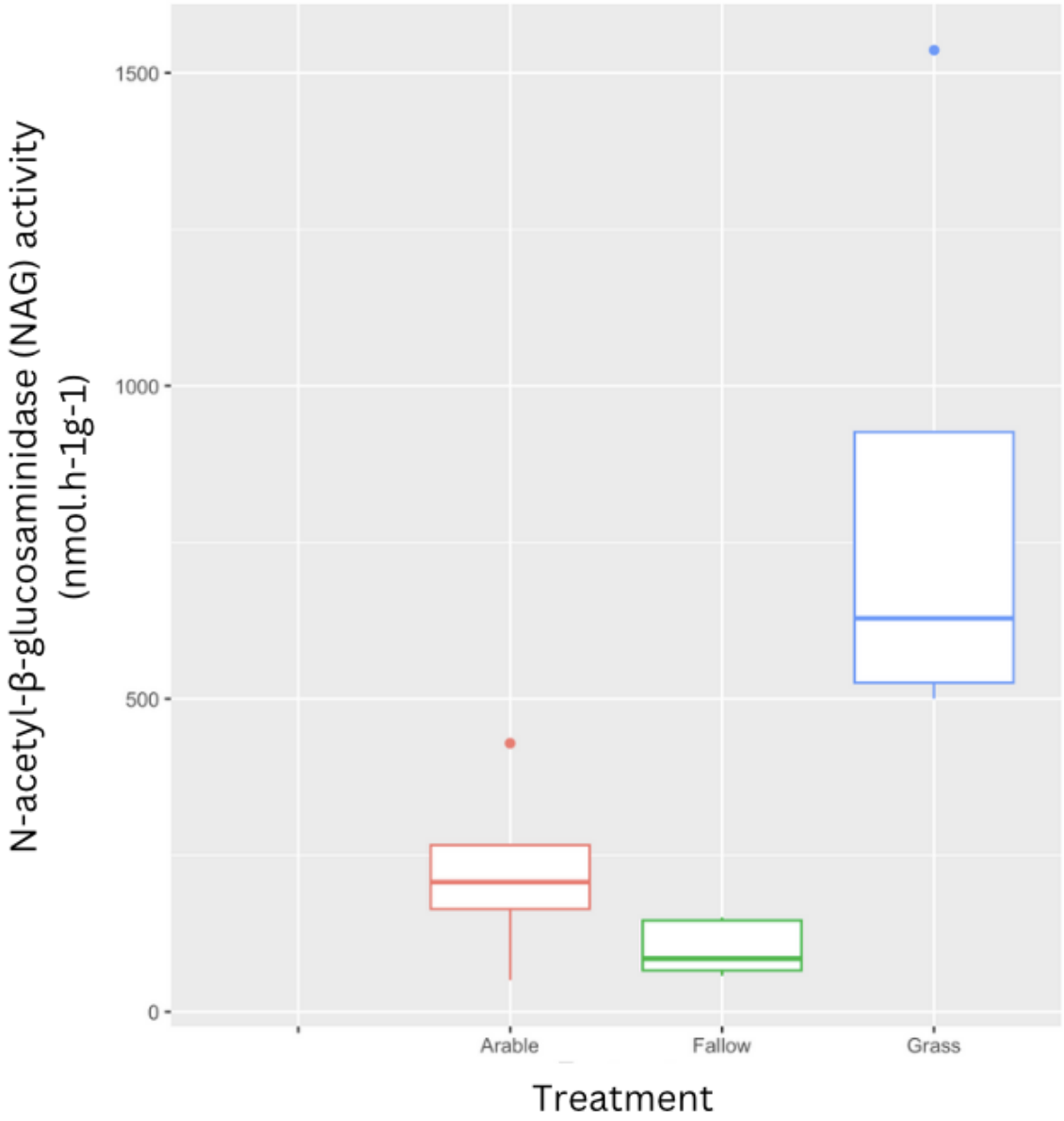


Figure 6: Boxplot showing the range of N-acetyl-β-glucosaminidase (NAG) activity (nmol.h^{-1}g^{-1}) under three different land management: Grass, Arable and Fallow treatments from the Highfield Ley-Arable long-term experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

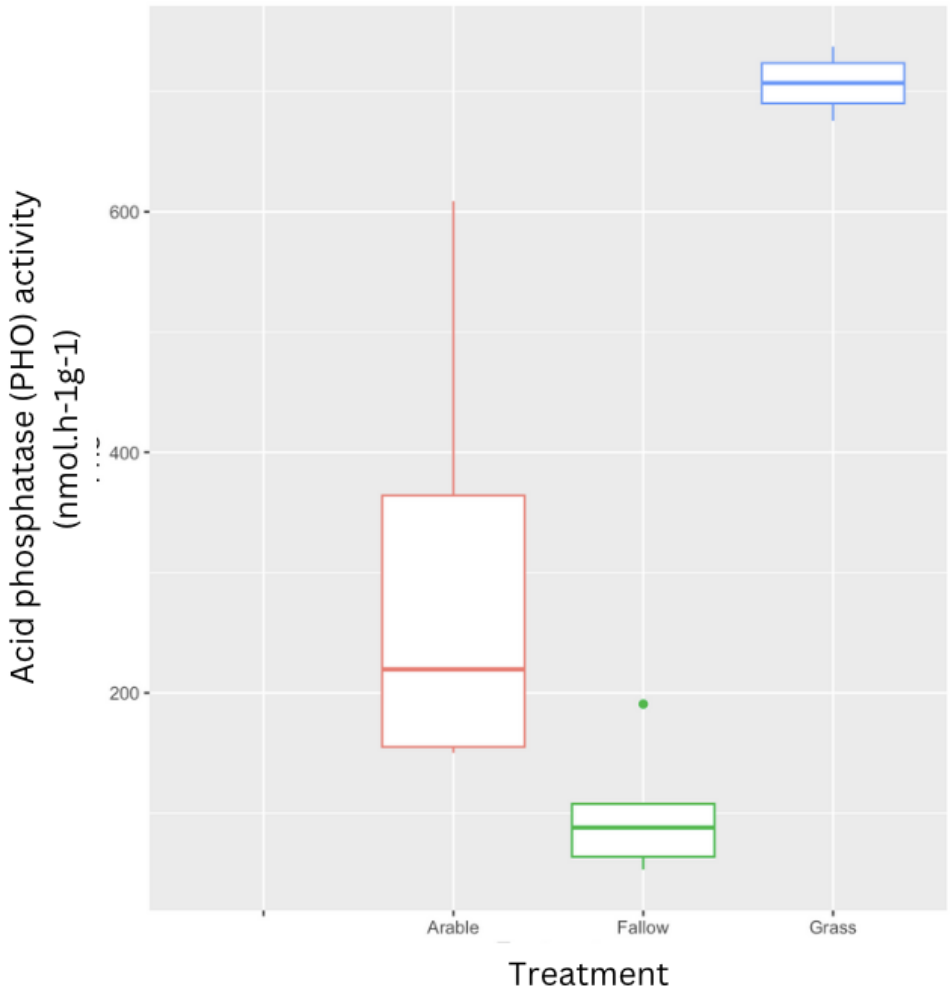


Figure 7: Acid phosphatase (PHO) activity (nmol.$h^{-1}g^{-1}$) under three different land management: Grass, Arable and Fallow treatments from the Highfield Ley-Arable long-term experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

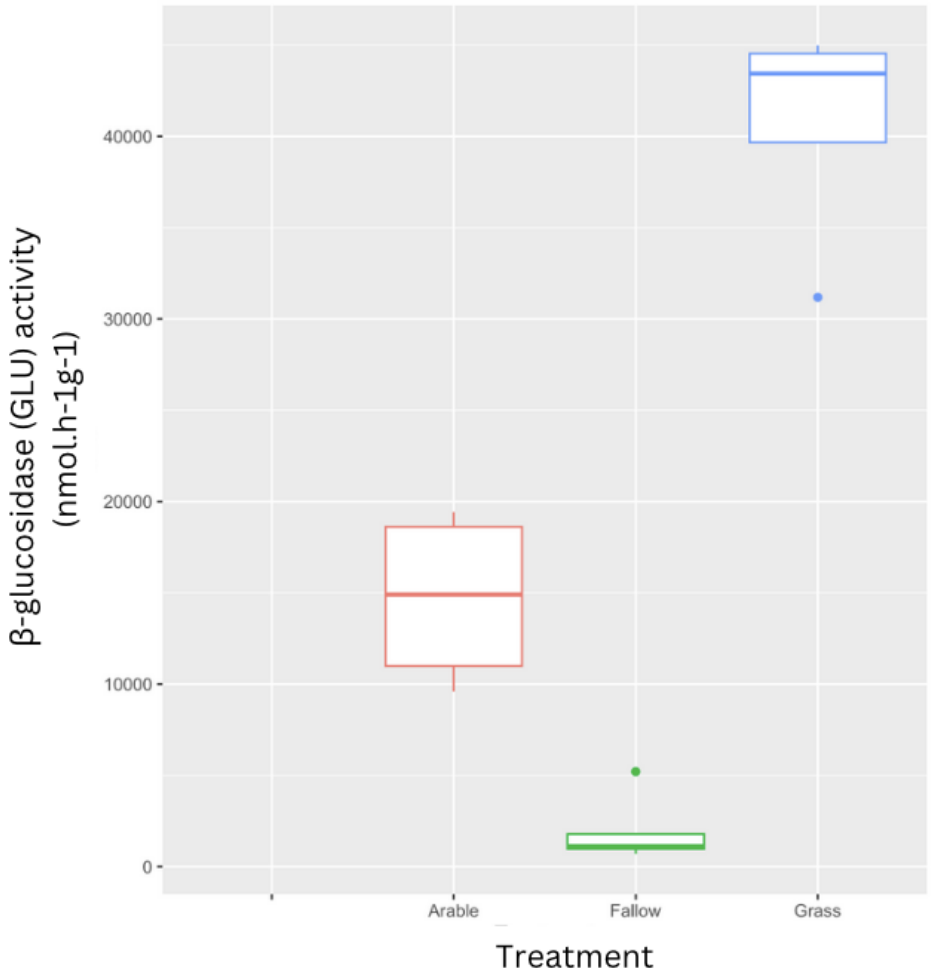


Figure 8: β-glucosidase (GLU) activity (nmol.$h^{-1}g^{-1}$) under three different land management: Grass, Arable and Fallow treatments from the Highfield Ley-Arable long-term experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

3.4.2 Organic Amendment Additions and EE Activity

The Woburn Organic Manuring experiment was used to understand if our chosen enzymes had different activities under certain organic amendment additions and understand how these were affected by other soil attributes. The results show that GLU, NAG, and PHO activities were significantly different under certain amendments. GLU activity was especially sensitive to amendment type and showed significant differences for the treatments: cover crop, FYM at 10 t ha^{-1} (DG10), and no additions ($p<0.000342$). In comparison with the other EE tested, GLU could distinguish between the highest number of amendments. PHO activity differed significantly only for treatments: FYM at 25 t ha^{-1} (DG25) and no additions ($p<0.00887$), whereas NAG activity could only statistically separate the grass ley treatment ($p<0.0215$).

We found that whilst GLU, PHO, and NAG could not separate all amendments, clear groupings could be seen in the activity data, as shown in Figures 9, 10, and 11. NAG, PHO and GLU all showed the smallest activity levels in soil with no additions (control). However, NAG showed the greatest activity in grass treatments, whilst PHO showed the greatest activity in the DG25 treatment. GLU showed the greatest activities in both the grass and the DG25 treatments.

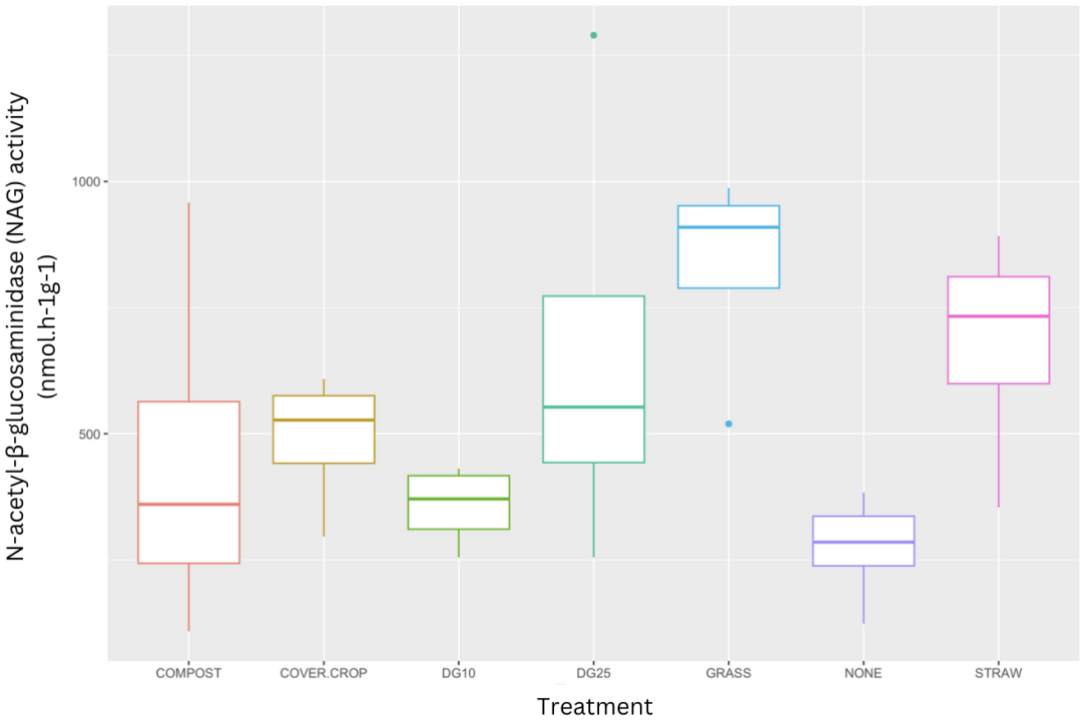


Figure 9: N-acetyl-β-glucosaminidase (NAG) activity ($nmol.h^{-1}g^{-1}$) under different organic amendment additions (compost, cover crop, farmyard manure under two concentrations (DG10 and DG25), grass, straw, and no additions) from the Woburn Organic Manuring experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

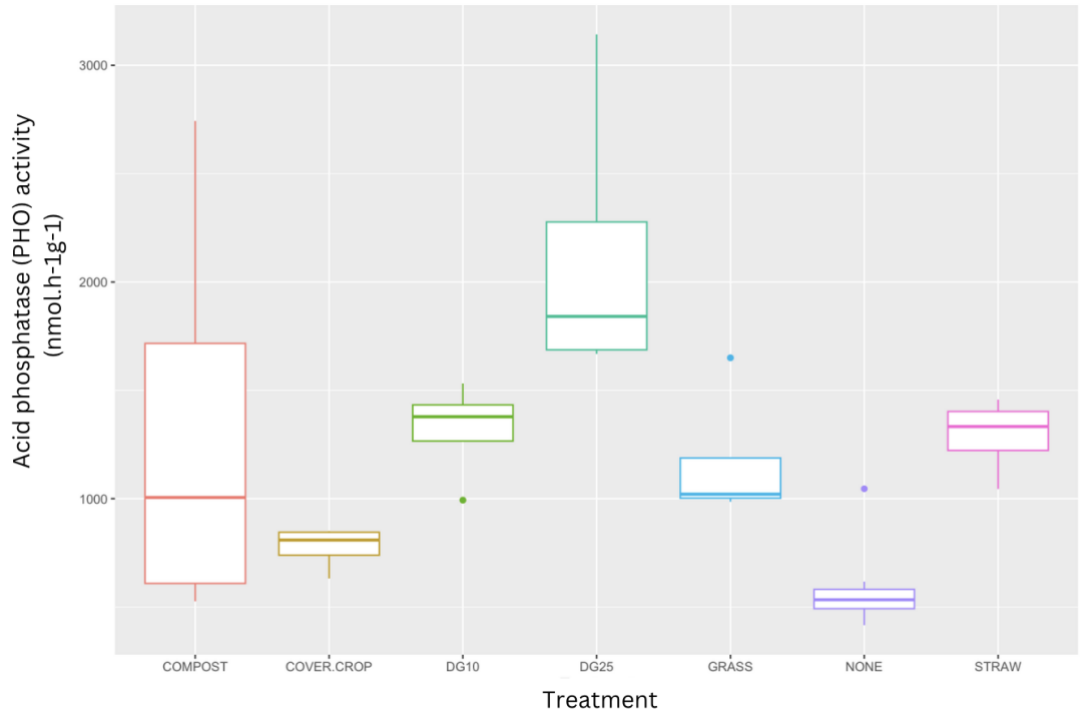


Figure 10: Acid phosphatase (PHO) activity (nmol.$h^{-1}g^{-1}$) under different organic amendment additions (compost, cover crop, farmyard manure under two concentrations (DG10 and DG25), grass, straw, and no additions) from the Woburn Organic Manuring experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

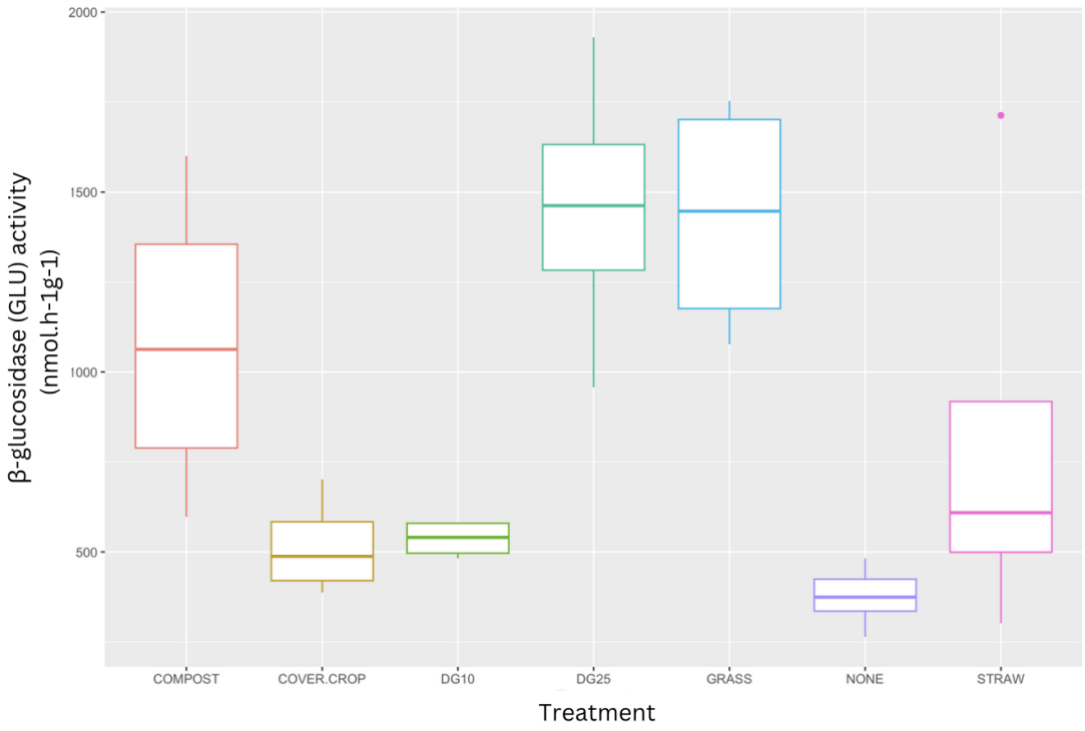


Figure 11: β-glucosidase (GLU) activity (nmol.$h^{-1}g^{-1}$) under different organic amendment additions (compost, cover crop, farmyard manure under two concentrations (DG10 and DG25), grass, straw, and no additions) from the Woburn Organic Manuring experiment. The box represents the interquartile range (IQR), with the median indicated by the line within the box. The whiskers represent the range of data outside the IQR, and outliers are plotted as individual points beyond the whiskers.

3.4.3 Soil Type and EE Activity

Another objective of this study was to see if enzyme activity was affected by soil texture. The soil at Highfield is a silty clay loam, whilst Woburn is a sandy loam. We only compared long-term grass and arable treatments at Highfield with arable and grass treatments with no organic additions at Woburn. Only GLU had significantly different activities under different soil textures, with greater activity levels under the silty clay loam (Figure 14). In contrast, PHO and NAG activities were not significantly different, as shown in Figures 12 and 13, respectively. Thus, our study indicates that GLU activity was the most significantly affected by soil type and texture.

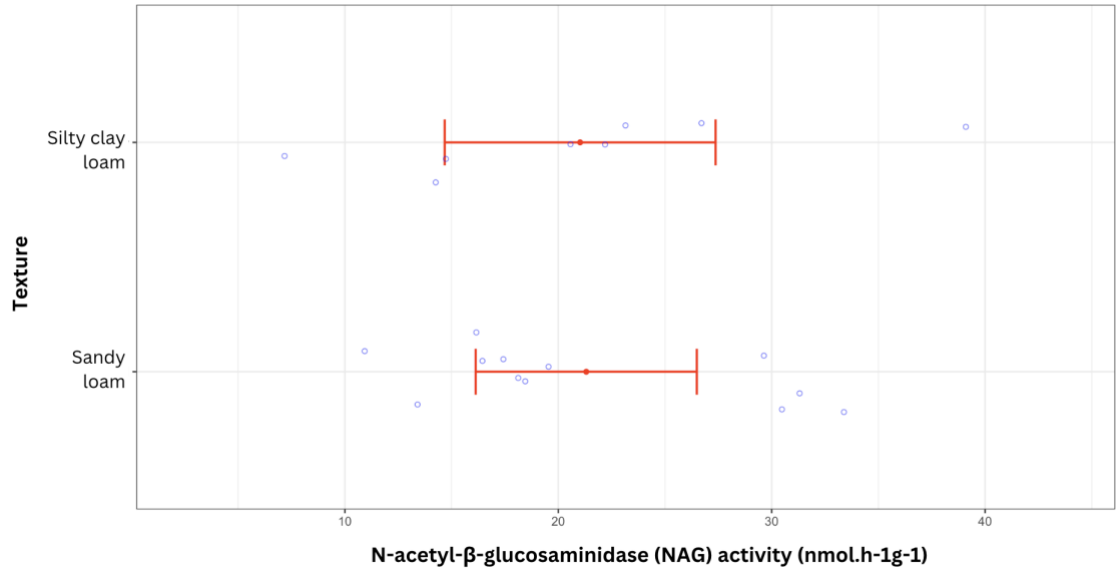


Figure 12: Predicted means with average LSD (5%) bar for N-acetyl-β-glucosaminidase (NAG) activity (nmol.$h^{-1}g^{-1}$) under different soil textures (silty clay loam and sandy loam) from Highfield Ley-Arable long-term experiment and Woburn Organic Manuring experiment. Overlap of the LSD bars shows that the differences between the means for NAG activity for each soil texture were not statistically significant.

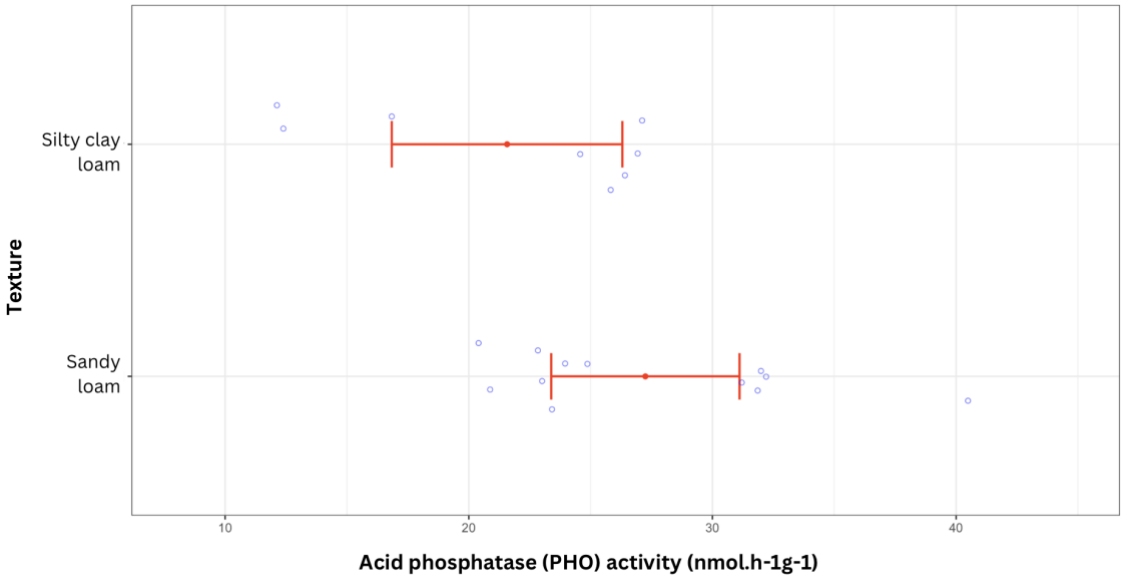


Figure 13: Predicted means with average LSD (5%) bar for acid phosphatase (PHO) activity (nmol.$h^{-1}g^{-1}$) under different soil textures (silty clay loam and sandy loam) from Highfield Ley-Arable long-term experiment and Woburn Organic Manuring experiment. Overlap of the LSD bars shows that the differences between the means for PHO activity for each soil texture were not statistically significant.

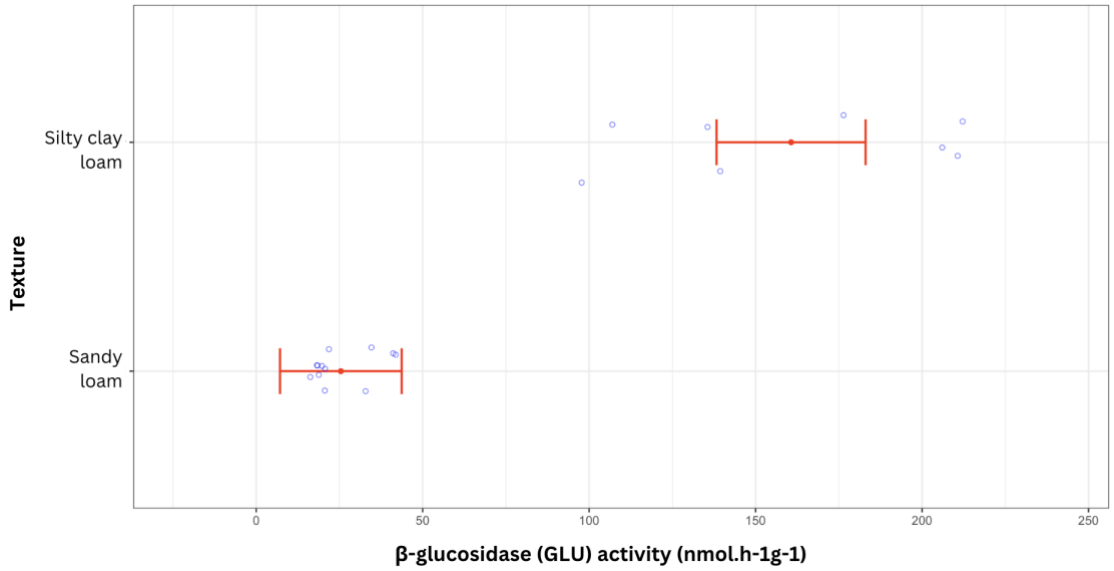


Figure 14: Predicted means with average LSD (5%) bar for β-glucosidase (GLU) activity (nmol.$h^{-1}g^{-1}$) under different soil textures (silty clay loam and sandy loam) from Highfield Ley-Arable long-term experiment and Woburn Organic Manuring experiment. No overlap of the LSD bars shows that the differences between the means for GLU activity for each soil texture were statistically significant.

3.4.4 Correlation Between EE Activity and Selected Soil Characteristics

To evaluate the ability of EEs to be used as a soil health indicator, we compared NAG, GLU, and PHO activities with important chemical indicators of the soil, such as organic carbon, total N, inorganic-P, and pH, under our full dataset from Highfield and Woburn experiments.

The results show that GLU activity was positively correlated with total C and total N but otherwise showed no or weak correlations with pH and inorganic P, respectively. Similar results were seen with NAG; however, NAG also showed a negative correlation with pH. In contrast, PHO activity was negatively correlated with inorganic-P levels but showed a positive correlation with pH. Overall, all three enzymes showed correlations with most chemical indicators, GLU showed stronger correlations with organic carbon and total nitrogen compared to NAG and PHO. Whilst NAG and PHO showed correlations with pH, only PHO was able to show a relationship with inorganic-P.

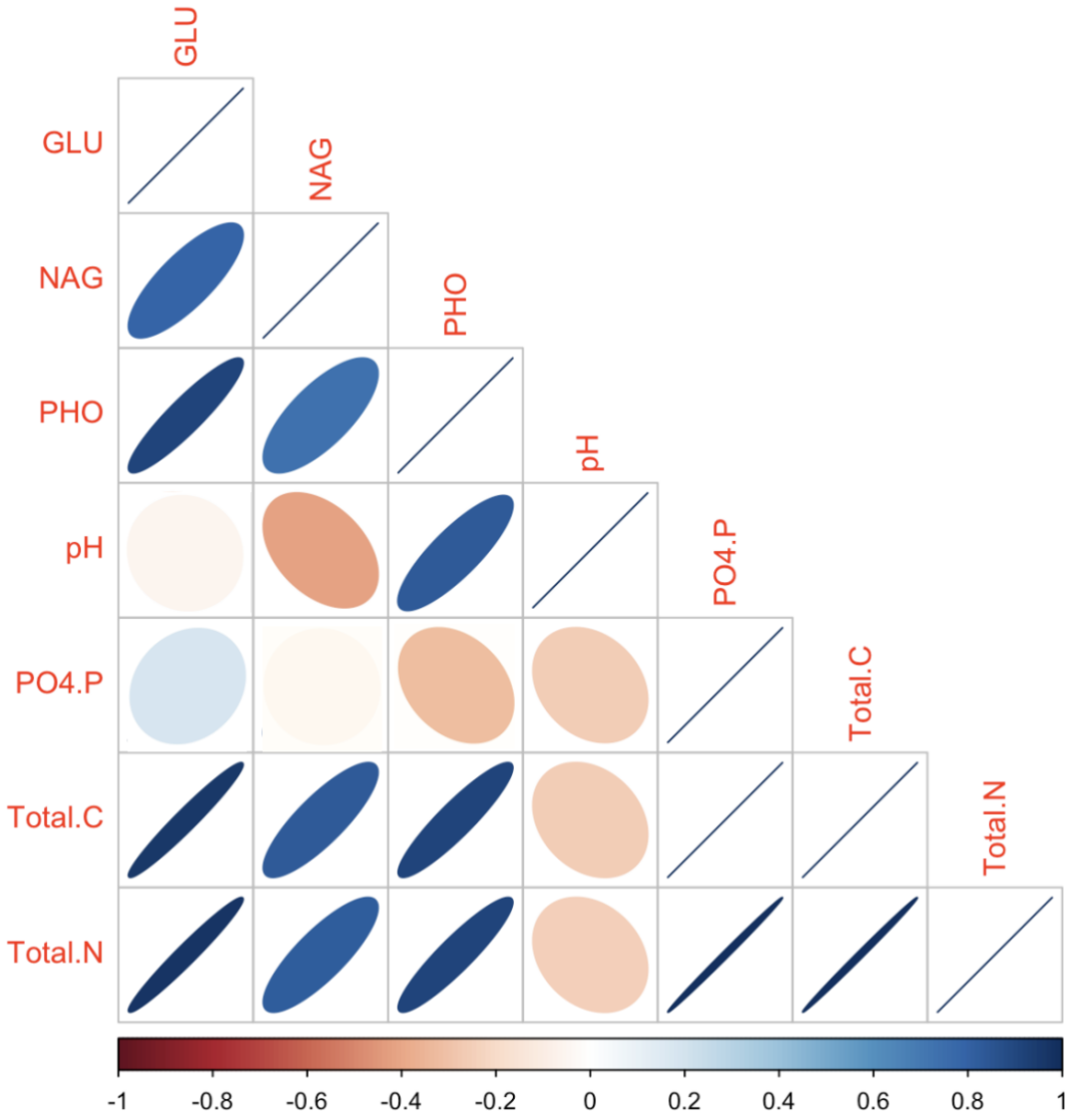


Figure 15: Correlation plot illustrating the pairwise correlations between enzyme activities for β-glucosidase (GLU), acid phosphatase (PHO), and N-acetyl-β-glucosaminidase (NAG), and chemical soil health indicators (pH, inorganic phosphorus (PO_4-P), Total Nitrogen (N), and Total Carbon (C)). The ellipses represent the strength and direction of the correlations, with darker shading indicating stronger correlations.

3.5 Discussion

3.5.1 Effect of Land Use on EE Activity

Farming practices can alter soil environments to facilitate plant growth by implementing conditions through the addition of fertiliser, pesticide, tillage, crop rotations and irrigation. Previous studies have shown how agricultural practices influence soil characteristics to affect nutrient turnover and microbial activity. These practices may affect soil enzymes and microbial activities differently (Curci et al., 1997; Schaeffer, 2021), and land use changes can affect soil ecosystems (Perring et al., 2015). In our study, land converted from grass to arable or fallow showed a significant change in EE activity. GLU, NAG, and PHO activity decreased under arable and fallow compared to grass treatments (Figures 6 to 8). However, we could not see a significant difference in NAG and PHO activities between arable and fallow treatments, and only GLU exhibited significantly different activities under all three treatments (Figure 8). Land use type alters the amount and composition of SOM, with permanent grass often showing considerably greater SOM concentrations (Haghighi et al., 2010). Therefore, it seems logical that we noted clear differences in enzyme activity between perennial (grass) and annually cultivated (arable and fallow) systems, with only GLU detecting the smaller differences between arable and fallow uses.

Disturbances in the soil can positively and negatively impact soil health and the environment. Arable and bare fallow treatments at the Highfield Ley-Arable experiment were ploughed. Whilst ploughing under arable lands can favour initial crop growth, it can also lead to OM losses and soil erosion (Kuipers. 1991). Moreover, ploughing affects the distribution of OM in the soil profile and, thereby, nutrient availability. As soil aggregates are broken and air is incorporated into the soil, labile SOM is exposed to microorganisms, which escalates the decomposition of SOM and crop residues. SOM decomposition releases nutrients, which in turn improves crop growth (Lv et al., 2023). Therefore, we could expect to see greater EE activity under arable treatments.

Overall, our results revealed that NAG, PHO and GLU had the greatest activity levels under grass, followed by arable, and then bare fallow, but only GLU had statistically different

activities under each treatment (Figure 8), which mirrors the general perception of soil health in these treatments (Neal et al., 2020). Ekenler et al. (2003) found that the most sensitive enzymes that reflect soil management practices were GLU and NAG. Similar findings were also reported by Madejon et al. (2007), who found that GLU activity levels were greatest under undisturbed systems. Furthermore, Pandey et al. (2014) showed greater GLU activities and increased levels of C and N, as well as microbial biomass with decreased soil disturbances. This was explained by increased substrate availability for microbial functioning and less soil disturbance, causing no further losses in OM (Monreal et al., 2000).

Green et al. (2007) noted greater PHO activity in undisturbed systems and suggested that PHO can detect changes in SOM under different soil management. Our results also showed greater PHO activity in undisturbed soils (Figure 7). Greater PHO activity in undisturbed treatments suggests better P cycling and soil structure. Organic P mineralisation, catalysed by PHO, provides P for plant uptake and growth. The ability of PHO, NAG, and GLU to detect changes in soil management and their strong interactions with SOM highlights their importance in acting as robust indicators of soil health.

3.5.2 Effect of Organic Amendments on EE Activity

Conventional agriculture can reduce soil fertility and SOM. Sustainable agricultural management practices must, therefore, be employed to ensure the future health of soils (Edmeades, 2003). The effects of organic amendments, including straw, compost, FYM, cover crops, and grass leys, in comparison with inorganic control treatment, on EE activity were examined. We found that compost and FYM treatments had the greatest PHO and GLU activity. We also discovered that NAG, PHO and GLU activity were all greater under FYM and compost additions (Figures 9 to 11). Similarly, Chang et al. (2007) found that PHO, NAG, GLU, and other biological factors were significantly greater in soils receiving compost compared to chemically fertilised soils. Compost and FYM improve soil health by increasing SOM and stimulating the chemical and biological properties of the soil (Saha et al., 2008). Further to this, Acosta-Martinez et al. (1999) found that GLU and NAG activities increased with leaf and N fertilisers.

Our results showed no significant changes in EE activity with straw additions, although NAG showed particularly great activity (Figure 9). However, Zheng et al. (2019) found that straw was the most effective amendment in increasing EE activity. NAG facilitates the breakdown of chitin, releasing N in the form of N-acetylglucosamine. This is then broken down into further N sources, such as ammonium and nitrate, which are essential nutrients for plant growth. Straw typically contains relatively low levels of N, and is generally C-rich rather than N-rich, which could explain the low levels of NAG activitity in our results.

Variations in OM inputs have been demonstrated to impact EE activity (Hernández et al., 2010). Hernández et al. (2010) observed greater NAG activity upon the addition of chitin as a result of OM additions, indicating that certain enzymes may be stimulated by the presence of the substrate they degrade. We also found that soils under cover crop and high OM see an increase in NAG activity.

Monokrousos et al. (2006) noted that changing from conventional to organic farming causes different responses in EE activities. In particular, the transitional period influences soil biological and chemical properties, where PHO was seen to increase under organic management over time. Most literature agrees that greater EE activities, particularly GLU and PHO, are seen under organic management systems, and chemical fertilisers cause a reduction in SOM and microbial activity in soils (Bandick et al., 1999; Pahalvi et al., 2021).

Furthermore, amendments such as FYM and compost showed great PHO activity (Figure 10), which can be attributed to the enrichment of SOM and microbial growth. Increased PHO activities are also likely due to enhanced P mineralisation following manure application (Sun et al., 2020). Additionally, Kalembasa et al. (2012) discovered that a combination of mineral N fertilisers and compost improved PHO activity, and the addition of P fertiliser to soils low in SOM also increased PHO activity. However, there was no significant PHO change when P fertiliser was added to soils of greater SOM. Thus, PHO synthesis and activity can be impacted by the addition of N and P fertilisers, particularly in soils of low SOM (Kalembasa et al., 2012; Saha et al., 2008). Moreover, Fließbach et al. (2007), who investigated SOM and biological soil quality indicators under organic and conventional farming, found that PHO

activities were greater in organically fertilised soils, suggesting that PHO was directly related to SOM content and was affected by farming practices. Thus, PHO activity can provide a good indication of SOM in soils.

3.5.3 Effect of Soil Characteristics on EE Activity

To serve as an effective indicator of soil health, EE must demonstrate evident relationships with soil functions and other indicators of soil health. We observed that PHO activity and total N had a positive correlation. However, our results found no correlation between PHO activity and total C (Figure 15). Kalembasa et al. (2012) studied the EE activity of soils after applying organic and mineral fertilisers. They also found that PHO activity increased relative to N amounts, suggesting that N fertilisers influenced PHO activity.

Furthermore, we noted a positive correlation between GLU and NAG activity with total C. This can be attributed to several factors related to microbial activity, SOM decomposition, and nutrient cycling. Total C in the soil is a key component of SOM, which serves as a substrate for microbial activity and provides a source of energy and nutrients for microorganisms. NAG activity is associated with the decomposition of chitin, which is often present in fungal biomass (Ekenler et al., 2003). The decomposition of chitin contributes to the release of C and N compounds, influencing the total C content and nutrient cycling in the soil, leading to a positive correlation between NAG activity and total C. We also recorded a positive correlation between NAG activity and total N. This finding was also consistent with Sotomayor-Ramírez et al. (2009), who recorded a positive correlation between NAG activity and total C and N values together with N mineralisation, suggesting that NAG could be a rate-limiting step in N mineralisation. Thus, NAG could be used to describe N cycling in the soil.

The positive correlation between GLU activity and total C in the soil is often observed and can be explained by the role of GLU in SOM decomposition and nutrient cycling. GLU hydrolyses glycosidic bonds, breaking down complex organic C-compounds into simpler sugars, contributing to the pool of soluble organic C in the soil. GLU activity is part of the broader spectrum of enzymes involved in nutrient cycling. As GLU breaks down SOM, it

releases not only C but also other nutrients such as N and P. This explains why we observed a positive correlation between GLU activity with total C and total N. Furthermore, whilst we did not find any correlation between GLU and inorganic-P, we found that PHO inhibition was correlated to greater PO_4 levels (Figure 15). According to Dick. (1997), P may inhibit PHO synthesis, and orthophosphate is a competitive inhibitor of PHO activity. This explains why our high PHO activity was negatively correlated with inorganic-P levels.

Soil pH influences PHO synthesis, release, and stability (Devau et al., 2009). Moreover, PHO typically exhibits optimal activity under acidic conditions, you would therefore expect that as the soil pH decreases, PHO activity would increase (Devau et al., 2009). However, we observed a positive correlation between pH and PHO activity. Soil pH can influence the decomposition of OM (Yan et al., 1996). Acidic soils may have greater SOM content, and the breakdown of organic P-compounds could contribute to increased PHO activity. The release of P from organic sources in acidic conditions may stimulate the activity of PHO, thus leading to a positive correlation (Gatiboni et al., 2018). Whilst we found no correlation between GLU activity and pH, Kim et al. (2021) discovered that as the soil pH increased from 4.5 to 8.5, GLU activity was seen to decrease. EE sensitivity to pH can aid in our ability to assess the effects of soil acidification, thus serving as a reliable biochemical indicator of soil health.

While our study highlights the potential for soil enzymes to be recognised as indicators of soil health due to their involvement in key biochemical processes, caution is warranted in making definitive conclusions about their practical use in soil health monitoring. The study presented here, conducted on a limited range of soil types and two long-term experiments, indicates that enzyme activity can reflect soil biological activity and management impacts. However, the extent to which these enzymes can serve as robust, practical indicators of soil health across broader soil types and conditions requires further investigation.

From a practical perspective, incorporating soil enzyme measurements into routine soil health assessments poses several challenges. Soil enzyme assays can be more time-consuming and technically demanding compared to traditional soil tests for physical and chemical properties, such as soil carbon and nitrogen. These traditional indicators are not only easier to measure but are also well-established proxies for soil health, particularly in

terms of nutrient cycling and organic matter content. The additional benefit of including enzyme activity in soil health assessments should be carefully considered, especially in terms of cost, complexity, and the incremental value they provide over more easily measured properties like soil C and N.

Given the limitations of the current study, further research is needed to validate the usefulness of soil enzymes as reliable indicators across a wider range of soil types and management systems. More comparative studies that assess the correlation between enzyme activity and other key soil health indicators would provide valuable insights. Additionally, investigating whether enzymes can offer unique information that other simpler soil properties do not capture will be crucial for their inclusion in practical soil health frameworks.

3.5.4 Effect of EE Activity Under Different Soil Textures

GLU was the only enzyme that had significantly different activities under different soil textures, with considerably greater activity under a silty clay loam (Figure 14). This shows that GLU activity was the most sensitive to soil type and texture. This could be explained by the differing nutrient and C levels in each soil. The ability of GLU activity to respond to different OA and soil types makes GLU an effective soil health indicator. Furthermore, GLU activity is integrated with SOM, microbial biomass, and C transformation and can provide early insights into the effects of agricultural management practices before other chemical and physical measurements. These features allow GLU to be used as an indicator of soil health. However, more knowledge is needed to show the intricate dynamics and factors influencing GLU activity. This will, in turn, enhance our understanding of agricultural management practices on soil fertility and soil health.

Whilst no single enzyme can describe the whole metabolic activity of the soil, certain soil enzymes are particularly important in SOM degradation. Of all the enzymes involved in cellulose hydrolysis, GLU is the most abundant in the soil. GLU, which catalyses cellulose degradation into glucose, plays a vital role in SOM decomposition. It provides an early indication of SOM turnover and can reflect the impact of changes in land management. Our

study found that GLU was the most sensitive to land use changes and correlated with more chemical soil health indicators compared to NAG and PHO. Furthermore, GLU is rarely substrate-limited and is easily detected. Due to this, we would recommend using GLU to aid in monitoring soil health. It is important to note that soil processes are complex, and the relationship between enzyme activity and soil properties can be influenced by various factors, including soil type, climate, land management practices, and the composition of the microbial community. Field studies and experiments, along with detailed soil analyses, should often be conducted to understand these relationships in specific ecosystems.

3.6 Conclusion

The use of EEs as a biological indicator of soil health has seen growing interest in recent years due to their significance in soil ecosystems. Research has previously shown the importance of EEs in nutrient cycling. Moreover, the existence and activity of EEs are heavily influenced by agricultural practices. Changes in the physical and chemical properties of the soil also provoke changes in EE activity; this shows that indicators of soil health are associated with each other. Organic applications, fertilisers, and land use affect nutrient cycling and microbial activity, as shown in our study. Significant variations in soil enzyme activities can be explained by the natural variability in environmental conditions and soil characteristics, which coincide with agricultural practices. Regardless, EE activity reveals changes in soil conditions and, therefore, soil health. However, site-specific evaluations need to be developed to accurately use EE activity as a soil health indicator as we identified differences between two soils of differing textures.

Our study found that EE activity was negatively correlated with soil disturbance, as EE activity was most significant under perennial grass, followed by annually cultivated arable and fallow treatments. Furthermore, EE activities provide valuable information on the effects of organic amendments in agroecosystems and can potentially predict SOM dynamics. Selecting specific enzymes as biological indicators for soil health depends on how they relate to crop yields and key soil parameters. Our study shows that GLU was associated with the most soil properties, followed by PHO and NAG, respectively. For this reason, we conclude that GLU activities would be the ideal candidate to describe soil health. However,

whilst soil enzymes are integral to soil health, for an accurate assessment, they need to be considered along with other physical, chemical, and biological indicators of soil health.

Moreover, the findings of this research suggest that soil enzymes hold potential as indicators of soil health, but conclusions must be tempered given the scope of this study. The evidence, drawn from a limited number of soil types and long-term experiments, highlights the need for further research to establish whether enzyme measurements provide significant added value beyond conventional soil properties like C and N, especially as our results show relatively strong correlations between soil EE activity and key chemical properties. Caution is therefore advised when proposing soil enzymes as standalone indicators, and their use should be considered complementary to traditional soil health metrics rather than as primary indicators.

4 Designing a Soil Health Index for Sustainable Agricultural Systems.

4.1 Abstract

Healthy soil acts as a reservoir and cycling system for nutrients essential for crop growth and hosts a diverse range of organisms, including bacteria, fungi, insects, and worms, contributing towards ecosystem stability and resilience. However, soil health cannot be directly measured. Instead, soil health assessments typically rely on a range of measurements from essential biological, physical, and chemical indicators. But due to the highly integrative nature of soil, it is difficult to develop general soil health indices.

Structural equation modelling (SEM) is one method to address the challenges associated with characterising soil health. It works by developing a structural model that outlines the relationships between different components of soil health, including physical, chemical, and biological indicators such as soil structure, nutrient levels, and microbial activity, and defining how each variable influences or is influenced by other variables. Path analysis within the SEM framework estimates the direct and indirect effects of variables on each other, helping to understand the causal relationships among different aspects of soil health. Model fit indices, such as the chi-square statistic, can then be used to assess how well the model aligns with the observed data.

Using Rothamsted Research's long-term experiments in the UK, we took measurements from essential soil health indicators with the aim of using SEM to design a soil health index that can describe soil health across different land management and soil types. Overall, we found SEM allows for a comprehensive understanding of observed indicators and latent variables that contribute to overall soil health. However, our proposed model proved to be unsuitable for application across all soil types and land management practices. Instead, its effectiveness was most apparent when applied to specific land uses and soil types; however, larger sample sizes are necessary to gain a comprehensive understanding. Nonetheless, based on our results, we believe soil scientists can leverage SEM to refine soil health assessment models and improve the accuracy of their measurements, as well as understand the effects of agricultural management practices on soil health.

4.2 Introduction

Healthy soils are crucial for ensuring food security across the globe. But soils are not just an important medium for growing crops; they are essential in providing important ecosystem services such as carbon sequestration, nutrient cycling, and water purification (Bunemann et al., 2018). This is supported by complex physical, chemical, and biological soil processes that all interact. Thus, a better understanding of soil components and the relationship between them should lead to a holistic approach to characterising soil functioning; such approaches have been used to define the health and quality of soils. Soil health is defined as "the capacity of a specific soil to function, within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and support human health and habitation" (Doran and Zeiss, 2000). Soil health can be affected by soil use and management approaches (Bunemann et al., 2018). Thus, knowledge of the effects of different soil management practices and land uses on soil function is needed in order to develop a working soil health index that accounts for these complex interactions (Morrow et al., 2016).

To develop indices to assess soil health, appropriate biological, chemical, and physical indicators and their interactions must be selected to provide a holistic overview of soil health (Haney, 2012). A multi-indicator index can comprehensively classify the full spectrum of soils and assess the effectiveness of land management approaches for degraded lands. Since the 1990s, major national soil assessment approaches have been implemented worldwide. Applied soil health tests were based on the relationship between the measured indicator and soil threats, functions, or ecosystem services; however, these relations were generally not quantitively tested (Bunemann et al., 2018). Furthermore, a review of two commonly used soil health tests in the US, the Cornell and Heaney soil health tests, concluded that neither test was able to correlate a soil health score with yields or differentiate between agricultural management approaches (Van ES et al., 2019). Overall, focusing on a particular soil threat, function, or ecosystem service does not provide enough information to indicate the holistic status of the soil. Thus, it seems that indices focusing on soil productivity over short growing periods do not capture the sustainability of soil systems in the long term.

4.2.1 Selection of Soil Health Indicators

The assessment of soil health involves three main steps: 1: Selection and measurement of soil health indicators; 2: Quantification of soil health indicators through direct measurement and assigning an appropriate score; 3: Integration of scored attributes to construct the final index by providing criteria for defining the weight of each indicator.

Selection of the most relevant indicators of soil health is a critical step, as they form the foundation for the soil health index (Cordoso et al., 2013). Correlation between indicators should be examined to minimise the number of measurements needed; this can be done using statistical tools such as Minimum Data Set (MDS) and Analysis of Variance (ANOVA) (Andrews et al., 2001). By doing this, we can exclude indicators that do not add additional or new information on the response to different crop and soil conditions.

Fine et al. (2017) suggest that indices based on a larger number of indicators (more than 5) are more informative. An index based on a smaller number of indicators was proposed by Haney (2012), who established a soil health index based on three biological attributes: soil respiration, dissolved organic nitrogen and dissolved organic carbon. Haney's index suggests that microbial activity and nutrient availability are the main factors describing soil health. A broader index developed at Cornell University included 39 different indicators, such as: physical (e.g. aggregate stability, penetration resistance, water holding capacity), chemical (pH, organic matter content, P, K) and biological parameters (soil pathogens, soil respiration) (Gugino et al., 2009). The additional soil health parameters were able to underline the overall status of soil and made the index more robust but were not able to provide more detail, unlike indices that used fewer parameters but focused on a particular soil attribute.

Moncada et al. (2014) assessed soil health by morphological classification using Visual Soil Assessment (VSA) and used decision trees based on predetermined thresholds. In contrast, Andrews et al. (2004) created a set of decision rules to select each indicator. The selection process was based on soil function, the management goals for each site, crop tolerance, climate conditions, and soil taxonomy, amongst other things. Each indicator could only be selected if it satisfied a unique combination of different criteria. Consequently, this decision

process was very complex and could be easily manipulated. Overall, decision-based frameworks used to select soil health indicators with many rules and criteria may increase the number of degrees of freedom for the statistical model used, thus weakening the relationship between soil properties and soil functions. Therefore, selecting a soil health indicator for the soil health index could be simplified by using statistical methods rather than a decision-based framework; this would also reduce the possibility of biases (Andrew et al., 2004).

Using an expert-based system, which is generally based on the accumulated knowledge of scholars, could be another tool when selecting soil health indicators. Ritz et al. (2009) used an experts-based framework to select biological soil health parameters for monitoring soil health on a national scale. Attributes selected by stakeholders and experts in the field should also adhere to current policy requirements and combine socio-economic and political factors associated with a specific soil health indicator. The expert-based approach assumes that a given expert understands the complexity of the mechanisms studied and that their knowledge can be translated accurately into the model. Thus, expert-based systems, at this selection stage, may differ when interpreting data gathered from different regions, land uses and points of view of the relevant experts. In contrast, Svoray et al. (2012) found that statistical models provided better predictive abilities compared to expert-based models. Moreover, complex expert-based systems may lack simplicity, which is crucial in allowing the application of minimum dataset selection, especially across a wide range of environments (Bunemann et al., 2018; Rinot et al., 2019).

4.2.2 Quantification of Soil Health Indicators via Direct Measurements and Scoring

Soil health indicators can be divided into quantitative or qualitative attributes (Amacher et al., 2007). The measured value of an indicator can be converted into a quantitative or qualitative unitless grade using calibration curves. These curves are generally based on a broad range of data with defined threshold values, and this provides appropriate scoring functions. The threshold can be defined from expert opinion, thresholds taken from literature, statistical models, or by considering observed values (Andrews et al., 2001).

Linear scoring functions are determined by dividing the measured value by the threshold value. The threshold value could be the minimum, maximum, or optimum value (Sharma et al., 2014). The score may be dependent on the variance of the specific soil health indicator; therefore, extreme outlier values may cause bias in the calculated scores. Moreover, linear scoring may not represent the current agronomic or environmental status of some attributes. Non-linear scoring usually assumes a normal distribution of the measured indicator and depends on non-linear patterns of response (Andrews et al., 2001). Andrews et al. (2001) identified that non-linear functions can better represent soil system attributes and thus are better suited for scoring functions compared to linear scoring functions. Furthermore, different functional transformations for each soil health indicator can be applied to different soil types, climate zones, and soil usages, allowing us to account for soil texture or other soil characteristics and functions when using non-linear scoring functions (Lilburne et al., 2004).

In contrast, when designing an SHI, some methods use raw values to construct the entire index and do not use a conversion step such as functional transformations, threshold values, or calibration curves. In Haney's test, the soil health index was developed using constructive functions. Also, the partial least squares (PLS) method used raw data for each soil health indicator to create a maximum explained variability of the model. According to Obade et al. (2016), PLS has great potential for predicting crop yields by selecting soil parameters, such as bulk density, soil organic carbon and electrical conductivity, and by considering the different soil types and soil management approaches. According to this approach, both qualitative and quantitative data can be utilised to construct the PLS model. When using this approach to construct the soil health index, no conversion step is required. However, this approach may be very challenging since a reliable definition and measurement of target values is essential.

4.2.3 Integration of Scored Attributes to Construct the Final Index

Current soil health indices typically rely on selected and representative soil attributes that specifically mirror a certain soil function or ecosystem service. Certain indices have

undergone testing to demonstrate their ability to see changes in soil functioning, while others have been validated against provisioning ecosystem services and productivity. These indices were discovered to overlook other crucial soil ecosystem services that are crucial for ensuring sustainable soil functioning for future generations.

Multivariate techniques such as principal component analysis (PCA) capture the maximum variance in the data with a smaller number of variables (principal components), thus providing a more concise representation of the original dataset. This reduction in dimensionality is useful for visualisation, noise reduction, and extracting meaningful patterns from the data. Whereas the PLS regression approach is adept at computing an index by considering the interrelationships among indicators and response variables like crop productivity (Vargas et al., 1999).

On the other hand, structural equation modelling (SEM) is a collection of multivariate statistical techniques employed to measure unobserved variables through sets of measured indicators by analysing the structural relationships between observed and unobserved variables (Maaz et al., 2023). Scientists have previously employed SEM to investigate intricate relationships in nutrient cycling and cropping systems in soils. However, its use in soil health indices has been relatively unused (Gama-Rodrigues et al., 2014). Compared to conventional methods used to assess soil health, SEM can simultaneously assess multiple soil health indicators and incorporate measurement errors in observed variables, accounting for measurement unreliability, thus ensuring a more accurate representation of the true underlying variables in the analysis. Furthermore, the use of a diverse range of soil health indicators serves to assess individual parameters and overall model fit, involving an examination of the disparities between the observed and model-implied covariance matrices, among other measures (Maaz et al., 2023).

In agricultural applications, SEM has been utilised primarily to examine the relationships between soil management practices, soil properties, and crop yield. Studies often focus on how variables, such as soil organic matter, nutrient availability, microbial activity, and physical structure influence soil health and productivity. SEM allows researchers to model these interactions, identifying pathways through which management practices directly or

indirectly affect soil properties and crop performance. However, SEM has not been widely utilised in soil science, unlike in the broader field of ecology, where it has been extensively applied to investigate how soil properties interact with other ecosystem components (Fan et al., 2016). For instance, SEM has been commonly used to identify the drivers of soil microbial communities, nutrient cycling, and organic matter decomposition. These factors are often influenced by both direct environmental drivers, such as temperature and moisture, and indirect factors, such as plant diversity. SEM enables researchers to disentangle these complex relationships by revealing how seemingly unrelated processes are interconnected, providing a more complete picture of ecosystem functioning.

As previous ecological literature demonstrates, SEM's ability to model direct and indirect drivers of soil properties makes it invaluable for addressing complex, system-wide questions in soil science and beyond (Eisenhauer et al., 2015). Future research in soil science could benefit from expanding SEM use to address non-agricultural ecosystems, particularly in the context of climate change, land degradation, and ecosystem restoration. By integrating data from diverse environmental variables and biotic interactions, SEM could play a crucial role in developing more comprehensive soil health indicators and improving our understanding of soil's role in maintaining ecosystem services.

Recent research has highlighted the potential of SEM as a valuable tool for understanding the complex interactions between soil properties, management practices, and ecosystem functions (Maaz et al., 2023). While SEM has been extensively applied in ecological studies to identify direct and indirect drivers of ecosystem properties, its use in soil science has been relatively limited. Maaz et al. (2023) further emphasise the importance of integrating SEM into soil health studies, particularly for exploring how soil properties interact with biological communities and external environmental factors. This study highlights the need to expand the use of SEM in soil research to provide a more nuanced understanding of soil health dynamics and ecosystem services.

The current global status of provisioning and regulating services reveals a substantial decline in key regulating services, including soil structure and water availability. (MEA, 2005). Nevertheless, over the last decade, there has been a notable surge in food production from

crops and livestock from agricultural intensification. Current literature on soil health assessments advocates for a more holistic approach that encompasses all parts of the soil ecosystem services. Moreover, in addition to provisioning services, the quantification of regulating and supporting services is essential.

The increased pressure on land resources, which subsequently results in accelerated land degradation coupled with reduced ecosystem services, highlights the urgent need to protect and maintain soil health. Soil serves as a multifunctional, dynamic, and complex ecosystem supporting three main ecosystem services: provisioning, regulating, and supporting services. A multivariate-complex soil health approach is needed whereby all three pillars of soil ecosystem services are quantitatively included in the assessment process of soil functioning. This approach will lead to the development of a new soil health index based on quantifying the relationship between soil health indicators and ecosystem services. Such an index would make a major contribution towards facilitating our understanding of the connection between the need for securing food for the world's growing population and the threat of expanding land degradation.

In this study, we aim to review and identify meaningful metrics for soil health that influence agricultural production and other ecosystem services. We also aim to use measurements of soil properties relevant to soil health, including the physical, chemical, and biological indicators of the soil under contrasting agricultural land uses and soil types in order to develop an integrative soil health index, allowing land managers to actively monitor the health of their soils.

4.3 Methods

We collected 96 samples across two sites and four experiments: Fosters Organic Amendment and Highfield Ley-Arable at Harpenden, UK and Woburn Ley-Arable and Woburn Organic Manuring Experiments at Woburn, UK (see Appendix A). These sites utilise several organic amendments and cover different land uses. Multiple samples were collected from the topsoil layer (0 to 23 cm depth) and homogenised for each plot per site. For the most part, samples were stored at 5 °C, except for the enzyme assays, where samples were

stored at -80 °C until sample analysis. Based on existing literature and expert-based opinions, we shortlisted several soil health indicators, which we split into different parameter groups (Table 1). Our chosen soil health indicators fall under soil chemical attributes: pH, total carbon (C), total nitrogen (N), inorganic phosphate (PO_4-P); physical: aggregate stability, bulk density, compaction; and biological: enzyme activity of N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU). Other indicators were also measured but were removed from our analysis because they were not complete and/or had great variability within plots.

4.3.1 Developing Soil Health Model Using SEM

SEM is a statistical tool that is increasingly applied in soil science to examine and model complex relationships among observed and latent variables. Here, we used SEM to quantify the latent variable “soil health” by examining the interactions and connections between different soil properties, environmental factors, and biological components. First, we identified key indicators that contribute to soil health, as detailed in Table 7. We hypothesised that environmental and management factors drive changes in biological, physical, and chemical soil properties. Together, these three key levers affect overall soil health. These hypotheses were represented as pathways in our model (Figure 16). We identified several key indicators of each of our hypothesised latent variables and selected the most appropriate methods for measuring each. We collected data through field measurements, laboratory analyses, or from existing studies.

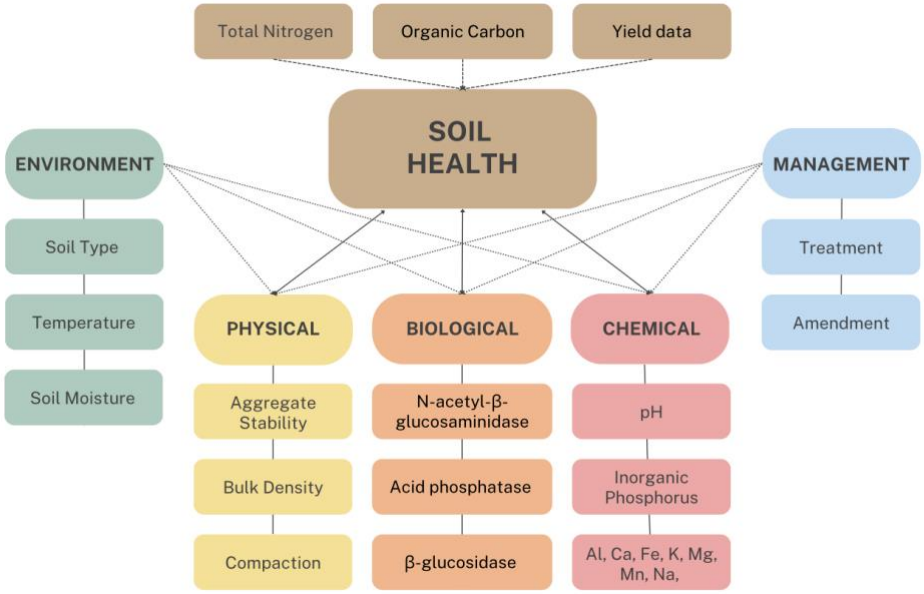


Figure 16: Structural equation meta-model of soil health. Latent variables are shown as rounded boxes, with their observed indicators listed in rectangular boxes of the same colour.

Understanding the SEM Components

Observed variables are the actual measurements we collected (e.g., Sand, Soil Moisture, Temperature, Organic Amendment, Grass, Arable, Chem1, Chem2, Bio1, Phys1, etc.). They are represented by square boxes in Figure 16. Latent variables are the unobserved variables inferred from observed variables. In our model, they represent broader constructs such as 'Environment', 'Management', 'Chemical', 'Biological', 'Physical', and 'Soil Health'. They are represented by rounded boxes. Loadings are the numbers associated with the arrows pointing from latent variables to observed variables. They show how strongly each observed variable is associated with its corresponding latent variable. Higher loadings indicate that the observed variable is more strongly related to the latent construct. Regression coefficients are the numbers on the arrows between latent variables. They show the strength and direction of the relationship between different latent variables. For example, a coefficient of -0.23* between 'Chemical' and 'Soil Health' indicates a weak but statistically significant negative relationship.

We scaled all numerical data and created dummy variables for categorical data. We fitted our model using the *sem* function of the {*lavaan*} package (Rosseel, 2012) in R (R Core Team,

2023). We report standardised coefficients to allow for simple comparison of different pathways. We assessed model fit with a chi-squared test of the hypothesis that observed and estimated covariance matrices are equal (P values < 0.05 indicate the model is a poor fit to the data).

Model Identification

Despite a large number of samples (n=96), our original proposed model was overidentified due to the complexity of the model (large number of parameters compared to the sample size, Portney 1988) and high multicollinearity (Variance Inflation Factor (VIF) >5) of our observed variables. To resolve these issues, we used Principal Component Analysis (PCA) to produce a smaller number of independent non-correlated variables as indicators for each of our physical, chemical, and biological latent variables (Table 7). We used the *prcomp* function in R (R core team) on our standardised and centred soil variables. The principal components identified in these analyses were then used as measurement variables in our SEM, replacing the observed indicators of biological, physical, and chemical (Table 8). As only one principal component was identified for physical, this was used directly in the regression models of our SEM in order to reduce parameterisation requirements and was not considered as part of a latent variable.

Table 7: Shortlisted soil health indicators summarised in each principal component used for the proposed soil health index.

Soil health indicator	**Principle Components**	**Latent Variable**
pH	Chem1, Chem2, Chem3	Chemical
PO_4-P	Chem1, Chem2, Chem3	Chemical
Major and minor trace elements (Al, Ca, Fe, K, Na, Mg, Mn)	Chem1, Chem2, Chem3	Chemical
β-glucosidase (GLU)	Bio1, Bio2	Biological
β-glucosaminidase (NAG)	Bio1, Bio2	Biological
Acid phosphatase (PHO)	Bio1, Bio2	Biological
Aggregate stability	Phys1	Physical
Penetration resistance	Phys1	Physical
Bulk density	Phys1	Physical

Table 8: Definitions of each latent variable used in our proposed model.

Latent variable	**Observed variable**
Soil Health	Yield + Total Carbon + Total Nitrogen
Biological	Bio1 + Bio2
Chemical	Chem1 + Chem2 + Chem 3
Physical (Phys1)	Phys1
Management	Treatment + Amendment
Environment	Soil Type + Temperature + Soil Moisture

Using SEM to Quantify Soil Health

We used global estimation to fit our SEM using Maximum Likelihood. This works to find the best global solution for our specified structural equations so that the covariance matrix of the solution best matches that of the observed data. Latent variables (soil health, biological, chemical, physical, management, and environment in our model) are constructs that do not exist in the data. They are defined by the indicators (observed variables) shown in Table 1. The model estimation adjusts the loadings for each indicator. Simultaneously, regression

coefficients are estimated for each structural equation (directional links between latent variables). Here, we report standardised coefficients and loadings to aid the interpretation of our soil health index. Standardised coefficients allow us to observe the relative strength and direction of relationships between variables of different units. Soil health scores ranged from +1 to -1.

4.3.2 Physical Soil Tests

Aggregate Stability

We used the fast-wetting component of the Le Bissonnais method to represent the mechanisms of aggregate disruption by slaking, and then measured the resulting aggregate size distribution (Le Bissonnais, 1996).

Soil samples collected from the field were hand-crumbled along existing pores and cracks over a single sieve stack with apertures of 5 mm and 3 mm. Soil aggregates collected in the lower 3 mm sieve (3-5 mm diameter) were then carefully transferred to a drying tin, removing any stones, and placed in the oven at 40 °C overnight. A subsample of 5 g of oven-dried aggregates were weighed and then transferred to a glass beaker filled with 50 mL of deionised water, leaving them to immerse. After 10 minutes, we carefully poured the water, trying to avoid losing any soil or further disruption. This step is not part of the standard Le Bissonnais method but follows the procedure used by colleagues at Rothamsted Research. Using a wash bottle containing methylated spirits (also known as denatured alcohol), we gently transferred the wet aggregates from the beaker to a 50 µm sieve submerged in a bowl of methylated spirits. The sieve was gently twisted ten times. The sieve containing > 50-µm stable aggregates were then removed from the methylated spirit bowl and left to air dry in the fume cabinet for 2 hours. Once air-dried, we brushed the > 50 µm aggregates onto a drying tin before placing the drying tin in the oven at 40 °C overnight.

The following day, we made a sieve stack comprising the following sieves from top to bottom: 2-mm, 1-mm, 500-µm, 200-µm, 100-µm, 50-µm, and receiver. We removed the dried > 50 µm aggregates from the oven, transferred the sample to the sieve stack and put a lid on. The sieve stack was gently shaken in a rotary motion for 30 seconds. We then

weighed the aggregates remaining in each sieve. Then, all > 50 µm aggregate fractions were placed into separate bottles containing sodium hexametaphosphate dispersing solution (2 g L^{-1}) to separate out the sand-sized (> 50-µm) primary particles (Kemper and Rosenau, 1986). After 4 hours of shaking, we assumed that all soil particles were fully dispersed. We thoroughly rinsed the remaining particles over a 50-µm sieve in a sink and transferred the > 50-µm sand-sized particles to a drying tin, which were then placed in the oven at 40 °C overnight.

After removing the dried > 50 µm sand particles from the oven, we tipped the sand particles into the same sieve stack described above and put a lid on before gently shaking for 30 seconds. We then weighed the sand particles retained on each sieve. All soil weights were determined to 2 decimal places (0.00 g). The stable aggregate size distribution was then calculated, which can be described by the conventional mean weight diameter (MWD) calculation, having accounted for sand-sized primary particles in each size class:

Equation 4:

$$MWD = \frac{\sum_{i=1}^{n}\left[\overline{\phi}_i.\left(a_{\phi_{i(a+p)}} - p_{\phi_i}\right)\right]}{a_d - \sum_{i=1}^{n} p_{\phi_i}}$$

where: MWD = mean weight diameter (mm)

$\overline{\phi}_i$ = mean aggregate diameter in size class *i* (mm)

$a_{\phi_{i(a+p)}}$ = oven-dry weight of soil (aggregates and sand particles) in size class *i* (g)

p_{ϕ_i} = oven-dry weight of sand particles in size class *i* (g)

a_d = total oven-dry weight of soil (aggregates and sand particles) (g)

Bulk Density

We collected intact soil cores from the desired depth increments using a soil core sampler (5×5 cm, diameter × height), ensuring that the soil core was undisturbed during collection to preserve its natural structure. The exterior of each soil core was cleaned to remove any adhering debris or organic material that may affect the accuracy of bulk density

measurements. The soil core was then extracted from the soil core sampler. The weight of each soil core before and after oven-drying at a temperature of 105°C was recorded. The difference in mass represents the moisture content of the soil core.

The bulk density of soil (Db, g/cm^3) is calculated from the mass of oven-dry soil (Md, g) and its field volume (V, cm^3), using the formula below:

Equation 5:

$$D_b = \frac{M_d}{V}$$

Total porosity (St, %) was calculated assuming a particle density of 2.65 g/cm^3 using the following equation.

Equation 6:

$$St(\%) = \left(1 - \frac{bulk\ density}{particle\ density}\right) \times 100$$

Penetration Resistance

Penetration resistance was measured using a penetrometer to assess soil density and strength. The penetrometer was positioned vertically over the soil surface at the sampling site, and the penetrometer probe was carefully inserted into the soil at a depth of 750 cm. The penetrometer probe recorded the resistance encountered as the probe penetrated the soil, taking measurements at every 1cm depth intervals throughout the soil profile. The average penetrometer resistance for each 10 cm depth increment was then calculated for each sample. Three repeat measurements were taken to confirm the results and address outliers. The penetrometer was calibrated according to manufacturer specifications.

Soil Moisture

The WET sensor (Delta-T Devices Ltd., Burwell, UK) measures three vital soil properties directly within the soil: water content, electrical conductivity (EC) and temperature. The WET sensor probe was calibrated to the manufacturer settings. Soil readings cannot be taken at any time throughout the day as readings will differ significantly if samples are taken during the morning vs. the afternoon as the soil dries as the temperature increases. As a result, we decided to take soil measurements during the morning.

To take soil moisture measurements, we pushed the WET Sensor into the soil. If the ground was hard or stony, we used an insertion tool to make guide holes first. Results appeared on the screen of a hand-held meter, which were then saved. The WET sensor was rinsed in tap water and wiped off after each use. Using the measuring tape, we took 12 readings evenly distributed throughout each plot. These values were then averaged to obtain one reading per plot.

4.3.4 Chemical Soil Tests

Total Carbon and Total Nitrogen

The LECO TruMac Combustion Analyser (LECO Corporation, St. Joseph, MI, USA) is a fully automatic instrument used for the determination of total N and C in soils and plant material, based on a modified version of the `Dumas' digestion method (Dumas, 1831).

Samples were weighed into ceramic `boats' or tin foil cups and placed on an auto-sampler. The sample then entered the combustion chamber, where the furnace and flow of oxygen gas caused the sample to combust. The combustion process converts any elemental C, S, and N to CO_2, SO_2, N_2, and NOx. These gasses are then passed through two anhydrone tubes to remove H_2O, a particle filter, and collected in a ballast tank. The gas was left to equilibrate before being released into an aliquot loop and through the infrared cells, where carbon was detected. Gas passed from the aliquot loop to the catalyst heater where NOx was reduced to N_2, then through Lecosorb to remove CO_2 and anhydrone to remove H_2O. The remaining

N_2 and helium carrier gas flowed through a thermal conductivity cell where the nitrogen was measured.

PO_4-P

Olsen P measurements as an indicator of soil P availability were determined in extractions from 5 g of air-dried, <2-mm soil with 0.5 M sodium bicarbonate at pH = 8.5. Soil samples were shaken for 30 min on an orbital shaker (120 rpm, 20 °C) and filtered through Whatman 42 filter paper. Phosphorus in the bicarbonate solution was determined by a phospho-molybdenum blue method on the Skalar SAN^{PLUS} System (Skalar Analytical B.V., Breda, The Netherlands), a continuous colourimetric flow analysis. Refer to Blitz et al. (1948) for a more detailed methodology.

pH

Soil pH is a measure of the hydrogen ion activity in soil solution (pH = $-\log_{10}[H^+]$). The electrometric pH reading is a product of complex electrode interactions between the electrode and the soil suspension; differences in soil:water extraction ratio, electrolyte concentration of the soil suspension, and spatial placement of the electrode can all affect this reading (Mclean 1982).

Two subsamples (15 g each) of a sieved (<2 mm) soil sample were placed into replicate centrifuge tubes. Tubes were capped to avoid moisture loss and, as necessary, stored in the refrigerator (5 °C) until ready for analysis within a day. Using a pipet dispenser or graduated cylinder, 30 mL deionised water was added to each tube to achieve a soil: water ratio of 1:2. Tubes were capped and shaken for a few seconds; the cap was then removed, and the slurry was allowed to equilibrate with atmospheric CO_2 and warm to room temperature for at least 30 minutes. The electrode of a pH meter, standardised at pH 7 and 4, was placed into the solution while gently swirling the slurry in the tube. pH was measured to the nearest 0.01. Between samples, the electrode was rinsed with deionised water.

Major and Minor Trace Elements

For analysis of major and trace elements, soil samples were digested in aqua regia (hydrochloric acid:nitric acid; 80:20 V/V) in open tube digestion blocks or digestion hot plates. The acids were removed by volatilisation, and the residue was dissolved in nitric acid (5% V/V) and filtered through a Whatman 40 filter paper. Primar (or equivalent) grade acids and 18 MΩ H_2O were used throughout.

Inductively Coupled Plasma - Optical Emission Spectroscopy (ICP-OES) was used to quantitatively determine trace elements. ICP-OES consists of a sample introduction system (autosampler, pump, nebuliser and spray chamber), a torch for plasma formation, transfer optics and an echelle polychromator with a solid state segmented array charged coupled detector (SCD). The sample introduction system brings the sample solution (in the form of an aerosol) to the plasma. The plasma is a highly ionised, very hot gas, which is stable and chemically inert with temperatures near 10,000 degrees K. The plasma excites the elements in the sample, causing them to emit light. Each element emits specific wavelengths of light and the light intensity corresponds to its concentration. Light passing through an entrance slit is dispersed by an echelle diffraction grating (selecting a wavelength) and then separated into visible and UV channels, which are detected by the SCD and converted to photoelectrons. Photoelectrons stored as electrical charge are transferred to the signal processing electronics, and the computerised data system converts the digital information into element concentrations.

All performance was strictly monitored using certified external standards alongside Rothamsted Research's in-house standard materials. Standards and check samples were monitored and recorded using Shewhart Control Graphs and computer-based quality control packages.

4.3.5 Biological Soil Tests

Enzyme Assays

We used an extract-based fluorimetric microplate assay with methylumbelliferone (MUB) as a fluorescence indicator to measure the activities of β-Glucosidase (GLU), N-acetyl-β-glucosaminidase (NAG) and acid phosphatase (PHO) using the substrates 4-MUB-β-D-Glucoside, 4-MUB-N-acetyl-β-glucosaminidase, and 4-MUB-phosphate, respectively (Criquet et al., 1999).

Table 9: Summary of the roles of soil extracellular enzymes that were used in our study.

Enzyme Type	**Enzyme Reaction**	**Enzyme name**	**Enzyme notation**	**Function of enzyme**
β-glucosidase	Cellobiose hydrolysis	β-glucosidase	GLU	C-cycling
Chitinase	Chitin hydrolysis	N-acetyl-β-glucosaminidase	NAG	C-cycling N-cycling
Phosphatase	Mineralises organic phosphate	Acid phosphatase	PHO	P-cycling

For the enzyme extraction, 3 g fresh soil was added to 50 mL of extraction solution (22.2 g $CaCl_2$, 20 g polyvinylpolypyrrolidone and 0.5 mL Tween 80 in 1 L dH_2O) and shaken at 150 oscillations min^{-1} for 1.5 hours. The samples were then centrifuged at 10,000 rpm for 10 minutes at 4 °C, and the supernatant was filtered through 1.2 µm filters (Whatmann™ GF/C) and dialysed for 12 hours using cellulose dialysis tubes (10-12 kDa cut-off) coated in polyethylene glycol (PEG). Following dialysis, the enzymes were recovered in 10 mL phosphate buffer (0.378 g Na_2HPO_4, 6.9 g KH_2PO_4 in 800 mL of dH_2O; pH 5.6), and the solution was separated into two 5 mL aliquots. One aliquot (live sample) was stored at 4 °C, and the second aliquot was pasteurised for 3 hours at 100 °C to deactivate enzymes, creating a positive control.

Assays were conducted in 96-well black microplates using 38 µL extract and 250 µL substrate (200 µM dilution). For each sample, eight wells (four analytical replicates each for live and control solutions) were filled, and the plates were incubated for 3 hours at 25 °C. Fluorescence was determined on a microplate reader (Fluostar Omega, BMG Labtech, Ortenberg, Germany) at 355 excitation, 460 emission and 975 gain.

A calibration curve was established from nine concentrations of MUB diluted in dH_2O (0-11.5 µmol L^{-1}, eight wells per concentration), read at the same settings. For each sample, the mean 'control' fluorescence was subtracted from the mean 'live' fluorescence and the resulting value (*F*) was then converted to activity (nmol h^{-1} g^{-1}) following Equation 4.

***Equation 7**:*

$$\text{Enzyme activity} = \frac{(F/V)}{c \times t \times M \times D}$$

Where *F* is the corrected mean fluorescence value for a given sample, *V* is the assay well volume (288 µL), *c* is the R^2 value of the calibration curve, *t* is the incubation time in hours, *M* is the sample mass in g dry weight, and *D* is the dilution coefficient 0.2 (50 mL initial extract volume concentrated to 10 mL final volume).

4.3.6 Yield Data

Farm staff collected crop samples at the time of harvest or cutting. Arable cereal crops were harvested using a Haldrup C-85 2m cereal plot combine (Haldrup GmbH, Ilshofen, Germany) with a 2-m cutting width across the full length of each plot. For grass crops, cuts were made with an Amazone Groundkeeper Smart Cut - GHS Drive 1500 flail mower collector (AMAZONE Ltd., Doncaster, UK; modified by Trials Equipment UK, Braintree, UK) with a 1.5-m cutting width across the full length of each plot. Two cuts of treatments under grass were made each year (mid-June and mid- to late-October). Dry matter content of grain or grass at harvest or cutting was calculated following oven-drying for 16 hours at 105 and 80 °C, respectively. Yields were then calculated as tonnes-per-hectare (Mg ha^{-1}) adjusted to 85% dry matter by convention for arable cereal crops, or as simply oven-dried basis for grass crops.

4.4 Results

Our main goal was to create a model that can generate a soil health score based on different land use histories and soil types, ensuring it can accurately capture the diverse conditions and variations observed across UK soils. Overall, most of our chosen soil health indicators (aggregate stability, bulk density, total organic carbon, total nitrogen, N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU), manganese (Mn), and inorganic phosphorous (PO4-P)) showed differences between soil type, with the exception of pH, penetration resistance, and some trace elements (Al, Ca, Fe, K, Mg, Na) (Table 10). Whilst NAG, PHO, GLU, aggregate stability, bulk density, penetration resistance, total organic carbon, total nitrogen, PO4-P, Ca, Na, and Mn showed significant differences between land management practices (Table 11), not all indicators were able to distinguish across all organic amendments.

Table 10: Statistics for each of our chosen soil health indicators across soil types (silty clay loam and sandy loam), including average values (mean) and standard deviation (sd), from all four long-term experiments (Highfield and Woburn Ley-arable and Fosters and Woburn Organic Amendment field experiments).

Soil type	**Aggregate stability***		**Bulk density***		**Penetration Resistance**	
	Mean	Sd	Mean	Sd	Mean	Sd
Silty lay loam	0.997	0.066	1.425	0.011	1366.05	371.846
Sandy loam	0.562	0.161	1.252	0.068	1873.89	234.626
	Total Organic Carbon*		**Total Nitrogen***		**Inorganic phosphorous***	
	Mean	Sd	Mean	Sd	Mean	Sd
Silty clay loam	1.550	0.126	0.153	0.010	29.075	6.757
Sandy loam	0.925	0.202	0.087	0.018	52.814	7.897

	β-glucosidase*		**N-acetyl-β-glucosaminidase***		**Acid phosphatase***	
	Mean	Sd	Mean	Sd	Mean	Sd
Silty clay loam	393.273	202.55	452.134	154.87	997.109	599.693
Sandy loam	1013.930	201.359	1158.55	337.41	3096.58	1135.012
	pH		**Aluminium**		**Calcium**	
	Mean	Sd	Mean	Sd	Mean	Sd
Silty clay loam	6.958	0.277	0.248	0.224	1835.53	332.867
Sandy loam	6.718	0.465	0.072	0.031	1420.53	394.25
	Iron		**Potassium**		**Magnesium**	
	Mean	Sd	Mean	Sd	Mean	Sd
Silty clay loam	0.067	0.038	201.900	77.483	76.917	16.554
Sandy loam	0.066	0.054	225.416	49.944	87.312	27.364
	Manganese*		**Sodium**			
	Mean	Sd	Mean	Sd		
Silty clay loam	20.582	10.664	8.881	4.967		
Sandy loam	4.443	2.941	3.971	1.666		

*Refers to soil health indicators that showed significant differences across soil types.

Table 11: Statistics for each of our chosen soil health indicators across land management (grass, arable and fallow), including average values (mean) and standard deviation (sd) from the Highfield Ley-arable field experiment).

Land management	**Aggregate stability***		**Bulk density***		**Penetration Resistance***	
	Mean	Sd	Mean	Sd	Mean	Sd
Grass	1.163	0.204	1.471	0.010	1974.984	299.376
Arable	0.977	0.036	1.457	0.008	1445.418	306.670
Fallow	0.841	0.041	1.359	0.011	932.276	177.097
	Total Organic Carbon*		**Total Nitrogen***		**Inorganic phosphorous***	
	Mean	Sd	Mean	Sd	Mean	Sd
Grass	3.742	0.328	0.322	0.023	74.840	6.560
Arable	1.648	0.074	0.161	0.006	32.960	1.479
Fallow	0.933	0.113	0.087	0.008	18.652	2.257
	β-glucosidase*		**N-acetyl-β-glucosaminidase***		**Acid phosphatase***	
	Mean	Sd	Mean	Sd	Mean	Sd
Grass	523.706	224.984	343.592	204.576	367.242	322.001
Arable	176.958	58.441	100.587	87.277	89.926	49.283
Fallow	241.863	22.322	31.725	17.943	29.678	13.774
	pH		**Aluminium**		**Calcium***	
	Mean	Sd	Mean	Sd	Mean	Sd
Grass	6.780	0.275	0.078	0.022	1606.381	615.630

Arable	6.763	0.343	0.122	0.039	1619.988	385.612
Fallow	6.926	0.217	0.115	0.035	1089.436	92.034
	Iron		**Potassium**		**Magnesium**	
	Mean	Sd	Mean	Sd	Mean	Sd
Grass	0.046	0.034	246.671	99.131	91.387	12.573
Arable	0.066	0.025	264.471	135.518	85.227	19.948
Fallow	0.101	0.059	187.672	21.838	75.632	11.065
	Manganese*		**Sodium***			
	Mean	Sd	Mean	Sd		
Grass	6.171	4.213	5.495	2.949		
Arable	10.483	6.226	6.858	3.553		
Fallow	2.287	0.681	2.796	0.407		

*Refers to soil health indicators that showed significant differences across one or more land management.

For our chemical, biological and physical soil health latent variables, we chose to use the first 3, 2 and 1 principal components as indicators, accounting for 88%, 72%, and 93% of the total variation with each latent variable, respectively. (See Appendix B for full PCA loadings and soil health indicator contributions for each principal component). Figure 17 shows our chemical, physical, and biological indicators, which have been transformed into a new coordinate system defined by the principal components, which are orthogonal axes that capture the maximum variance in the data. The position of each data point in the 3D space is determined by its scores on the principal components. The original soil properties are projected onto the new coordinate system and clusters of arrows in the plot indicate similarities among the soil health indicators.

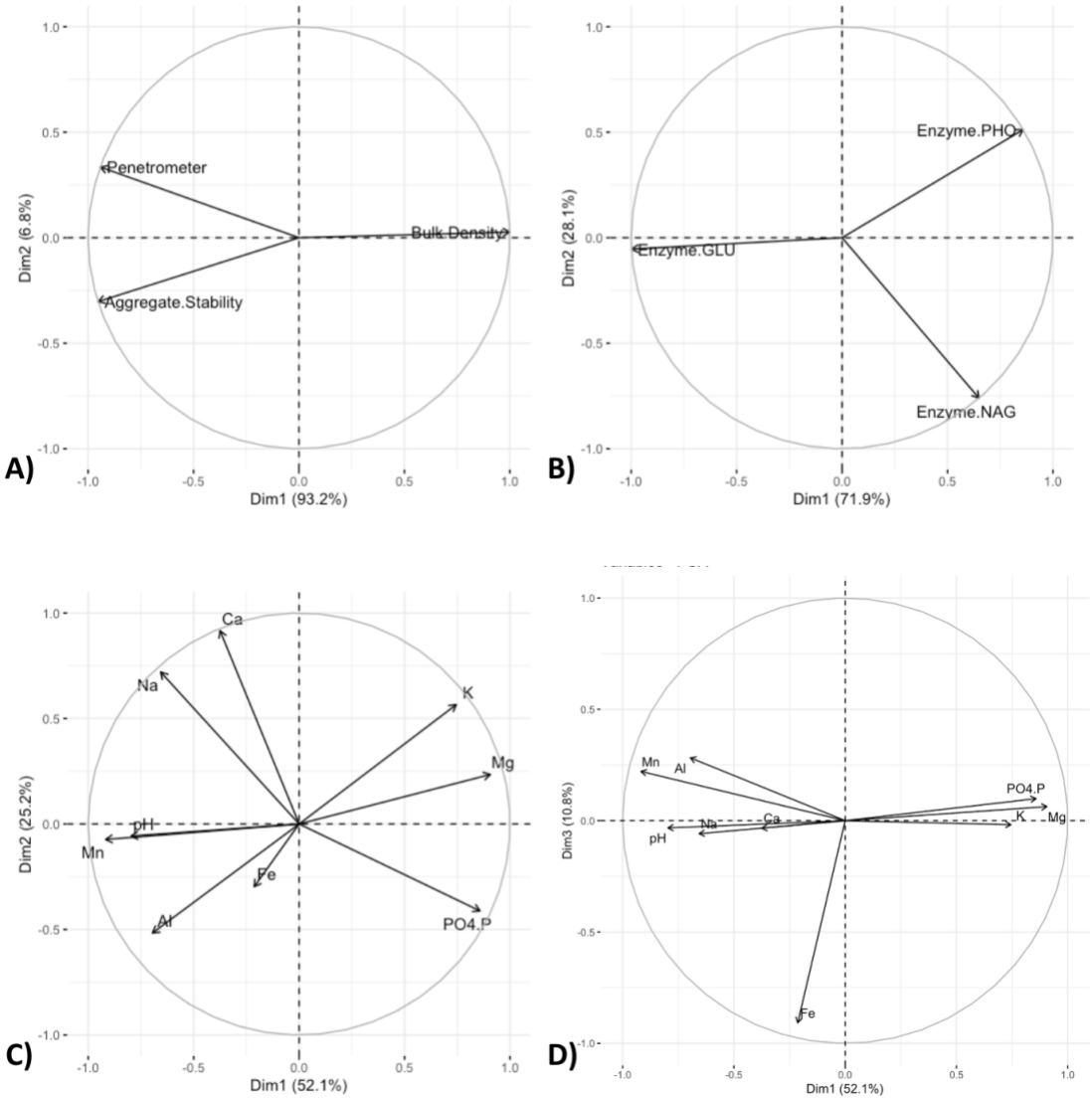


Figure 17: Three-dimensional (3D) Principal Component Analysis (PCA) plot representing the distribution and relationships among our A) physical, B) biological, C and D) chemical data points (C shows dimensions 1 and 2, and D shows 1 and 3 for chemical data points). The loadings for each soil property are shown with a solid arrow. The length of the arrow shows the size of the contribution to each principal component.

Our resulting SEM model contained a two-tier structure. The first tier included the measured physical, chemical, and biological soil health indicators as described in Tables 7 and 8, and the second-tier factor consisted of the 'overall soil health', which had the indicators total organic carbon, total nitrogen, and yield collected from arable and grass plots (Figure 18). The model additionally incorporated regressions to account for the influences of environment and management on each soil health indicator, as discussed earlier. To ensure

the model's accuracy, we conducted post-hoc adjustments to introduce and independently estimate parameters governing the covariance and correlation among specific soil health indicators. However, overall, our model fit was poor, with a CFI (Comparative Fit Index) of 0.914, RMSEA (Root Mean Square Error of Approximation) of 0.183, SRMR (Standardised Root Mean Square Residual) of 0.116, and a significant Chi-square result (Chi-sq = 0.00, P≤0.00, df = 81), suggesting that the model should be developed further. The significant chi-squared value suggests that there were discrepancies between the covariance matrices of the proposed model and the observed data.

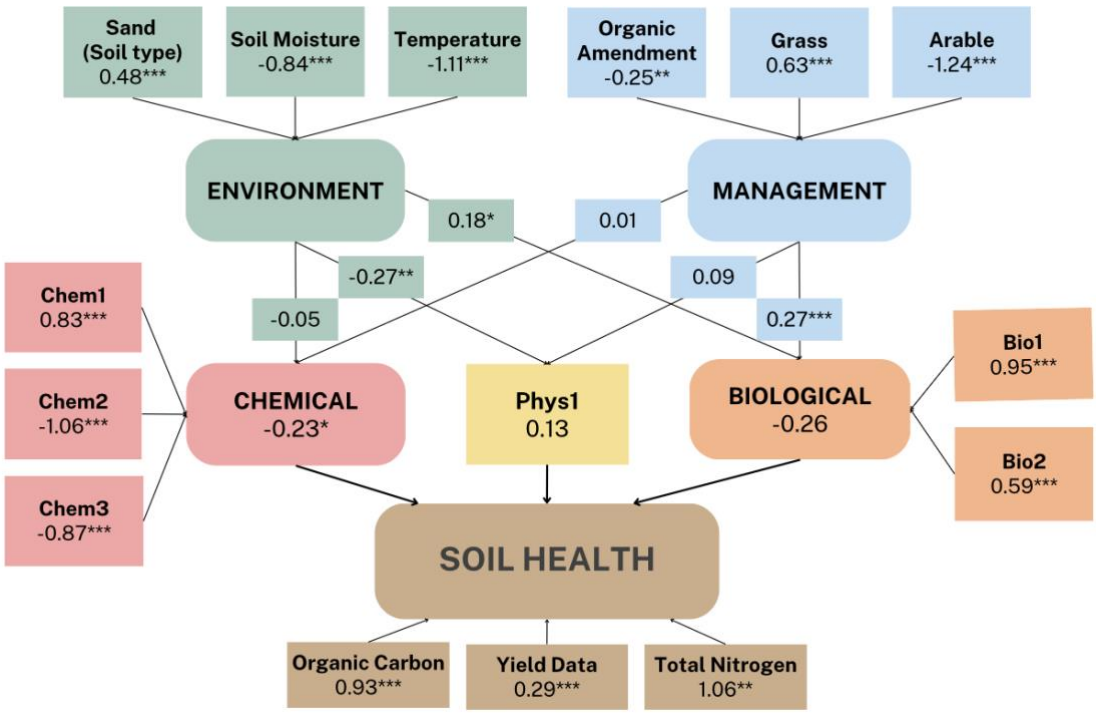


Figure 18: Proposed model showing how biological, chemical, and physical factors contribute to soil health. The model includes how the environment (soil type, temperature, and soil moisture) and management (treatment and amendment) can influence soil health. Soil health was defined using total nitrogen, carbon, and yield data collected from each plot. Numbers in each observed indicator box linked to latent variables depict the loading of each indicator. Numbers in each latent variable box linked to other latent variables show regression coefficients. Observed indicators are represented in square boxes, and rounded boxes are latent variables. Significant effects are indicated *p<0.05, **p<0.01, ***p<0.001.

The SEM presented in this study was designed to assess the relationships between various environmental, management, and soil health indicators. The model shown in Figure 18 integrates both latent and observed variables to determine how environmental factors (soil type, temperature, and soil moisture) and management practices (such as organic amendments and land use types) contribute to overall soil health. Soil health in this model is quantified using three key indicators: total nitrogen, organic carbon, and yield data.

The environmental latent variable in our SEM was influenced by three key indicators. Sand content had a positive loading on the environmental variable (0.48***), indicating that soils with higher sand content tended to score higher on this environmental axis. Soil moisture had a significant negative loading (-0.84***), suggesting that drier soils are associated with higher environmental scores in the context of this model. Temperature also had a significant negative loading (-1.11***), indicating that cooler temperatures are associated with higher environmental scores.

The environmental variable positively influenced the biological component of soil health (0.18*), suggesting that favourable environmental conditions (as defined by the SEM) lead to higher biological activity in the soil. Conversely, environmental conditions had a negative effect on the physical component (-0.27**), implying that these conditions might be associated with lower levels of key physical indicators such as aggregate stability and bulk density. Although not statistically significant, the environmental variable showed a weak negative relationship with the chemical component (-0.05).

The use of organic amendments had a slight negative loading on the management latent variable (-0.25**), suggesting that sites receiving organic amendments might score lower on the management axis, potentially due to the varied effects of amendments across different soil types. Grass cover had a strong positive loading (0.63***), indicating that pasture or grassland management practices contribute positively to the management score. Whilst arable land use had a significant negative loading (-1.24***), suggesting that cultivated land negatively impacts the management score.

The management variable positively influenced the biological component (0.27***), indicating that better management practices (higher management scores) lead to increased biological activity, as reflected by enzymes GLU and NAG. The management variable showed a negligible, non-significant positive influence on the chemical component (0.01), implying a weak or minimal effect of management on chemical indicators. Similarly, there was a small positive effect of management (0.09) on the physical component, suggesting that better-managed soils might have slightly improved physical properties like aggregate stability and bulk density.

Total nitrogen had a strong positive loading on soil health (1.06***), indicating that higher nitrogen levels are a critical indicator of better soil health. Organic carbon also had a strong positive loading (0.93***), further emphasising its importance as a key component of healthy soils. Yield data, while still positively correlated with soil health (0.29***), had a lower loading compared to nitrogen and carbon, suggesting that while crop yield is an important measure, it may be influenced by multiple factors beyond those mentioned in our SEM alone.

The indicators for the biological, physical, and chemical latent variables were built from our PCA (Figure 17; Appendix B). The biological component is significantly influenced by Bio1 (0.95***) and Bio2 (0.59***), with Bio1 distinguishing between soils with varying levels of GLU activity and Bio2 reflecting differences in NAG and PHO activities. It follows that sites with greater scores on our biological latent variable generally exhibited higher GLU and PHO activity, with NAG activity contributing variably depending on the balance between Bio 1 and Bio2 (Appendix B, Table B4).

The chemical component is strongly influenced by Chem1, which showed a positive effect (0.83***), while Chem2 and Chem3 had negative loadings (-1.06*** and -0.87***, respectively), indicating a complex relationship between the chemical indicators and soil health. Meanwhile, the physical component, represented by variables like compaction, bulk density, and aggregate stability, had a positive but non-significant loading on soil health (0.13), highlighting its importance but also the need for further investigation into its role.

Together, these three components (Biological, Chemical and Phys1) determined the final soil health variable. Soils with great physical values (0.13), together with small biological and chemical values (-0.26 and -0.23*, respectively), gave large scores for soil health; these soils typically had large C, N and Yield as these indicators all contributed positively to the latent variable.

Summary of Key Relationships in our Model

Environmental Effects

Sand (0.48***), Soil Moisture (-0.84***), and Temperature (-1.11***) have significant loadings on the 'Environment' latent variable. This suggests that these factors are key components of the environmental conditions affecting soil health. The 'Environment' variable has a small positive influence on 'Chemical' properties (0.18*) but a negative impact on 'Biological' properties (-0.27**), indicating that environmental conditions might negatively affect biological soil health but have a minor positive effect on chemical properties.

Management Effects

Grass (0.63***) and Arable (-1.24***) significantly contribute to the 'Management' latent variable, indicating that these land uses strongly influence management practices. 'Management' has a positive relationship with 'Biological' properties (0.27***), suggesting that certain management practices improve biological soil health.

Chemical, Biological, and Physical Impacts on Soil Health

'Chemical' properties have a small but significant negative effect on overall 'Soil Health' (-0.23*). This could indicate that certain chemical properties (like excess nutrients or contaminants) might degrade soil health. 'Biological' properties also negatively affect 'Soil Health' (-0.26), which could reflect issues like nutrient competition or the presence of certain enzymes in unhealthy soil conditions. 'Physical' properties show a positive but non-significant effect on 'Soil Health' (0.13), suggesting that while physical structure is important, its impact may be less direct or less pronounced compared to chemical and biological factors.

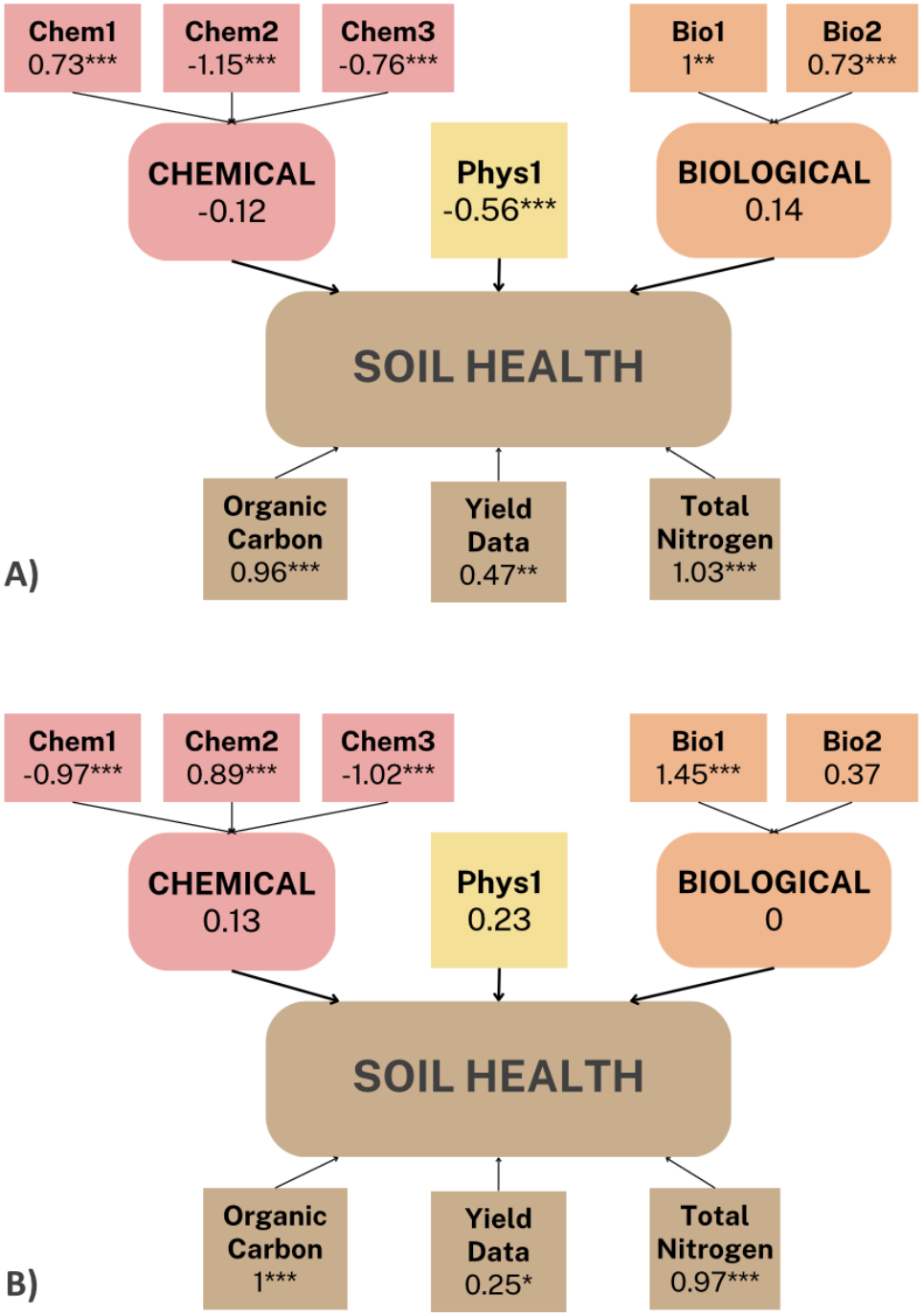


Figure 19: Modified model only using collected soil data from A) Clay or B) Sandy soil types. Numbers on arrows depict the strength of correlation each variable has in describing soil health. Physical soil parameters under clay and sandy soils have an inverse effect, -0.56*** and 0.23, respectively. Significant effects are indicated *p<0.05, **p<0.01, ***p<0.001.

From our preliminary analyses of our various soil properties and our understanding of the way in which they interact, we hypothesised that the influence of different soil types, amendments, and treatments may interact with the various physical, biological, and chemical properties. Due to the limited sample size, we were unable to fit an SEM with these interaction terms included. Instead, we subsetted our data collected from clay and sandy soils and fitted additional SEMs, omitting the management and environment variables.

Figure 19 shows differences between the resultant SEMs, particularly from our physical soil health indicators, where we see that physical parameters under clay and sandy soils had an inverse effect of -0.56 and 0.23, respectively. Similarly, we saw differences across our model when we only input collected soil data from arable or grass treatments.

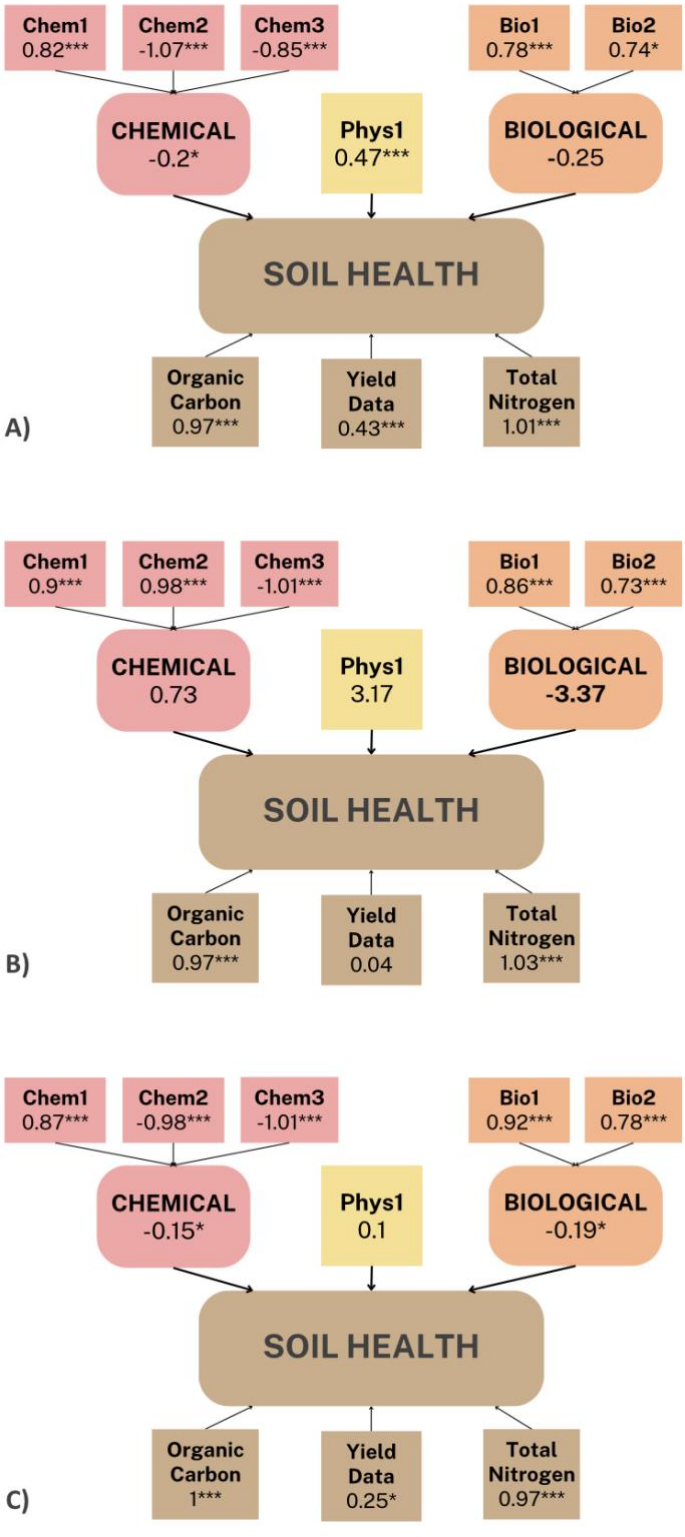
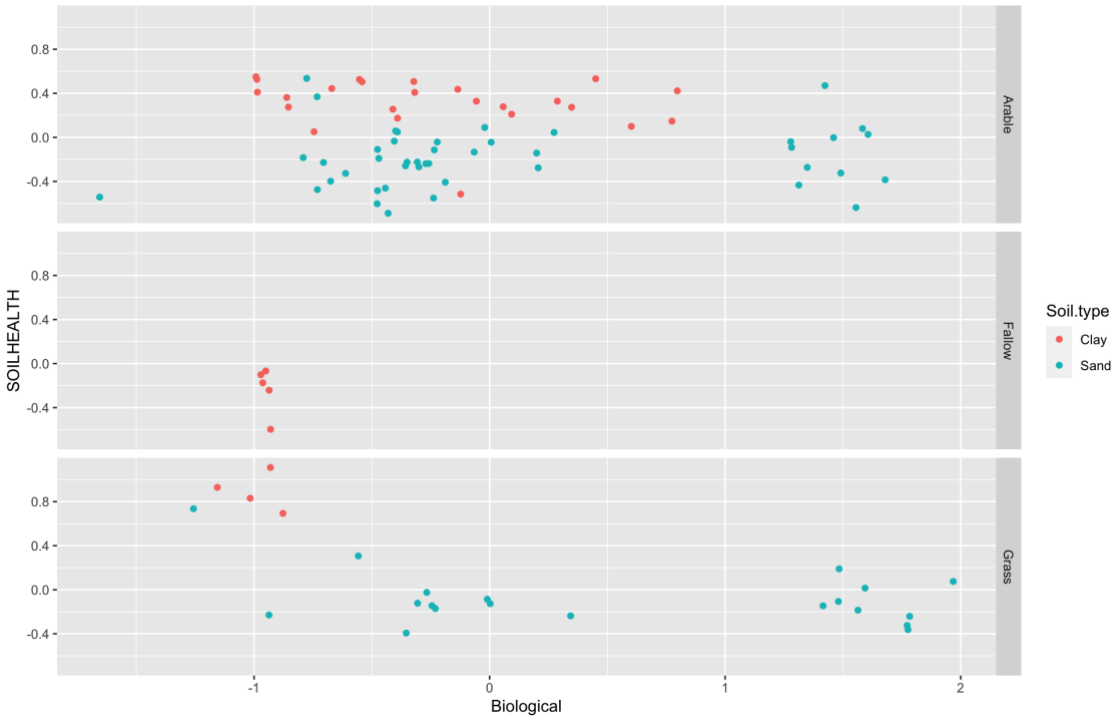
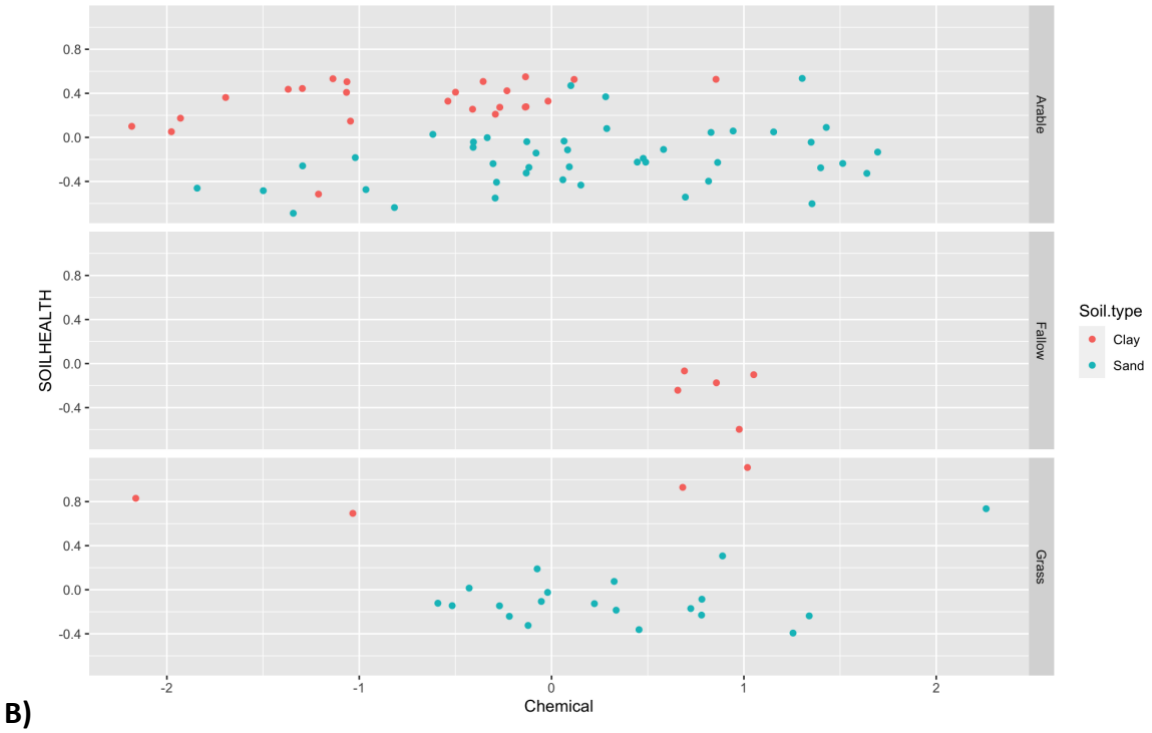


Figure 20: Modified model only using soil data collected from A) Arable, B) Grass, or C) No added amendment data. Numbers on arrows depict the strength of correlation each variable has in describing soil health. Chemical soil parameters under arable, grass, and soils with no added amendments had an inverse effect, -0.2* and 0.73, -0.15*, respectively. Significant effects are indicated *p<0.05, **p<0.01, ***p<0.001.

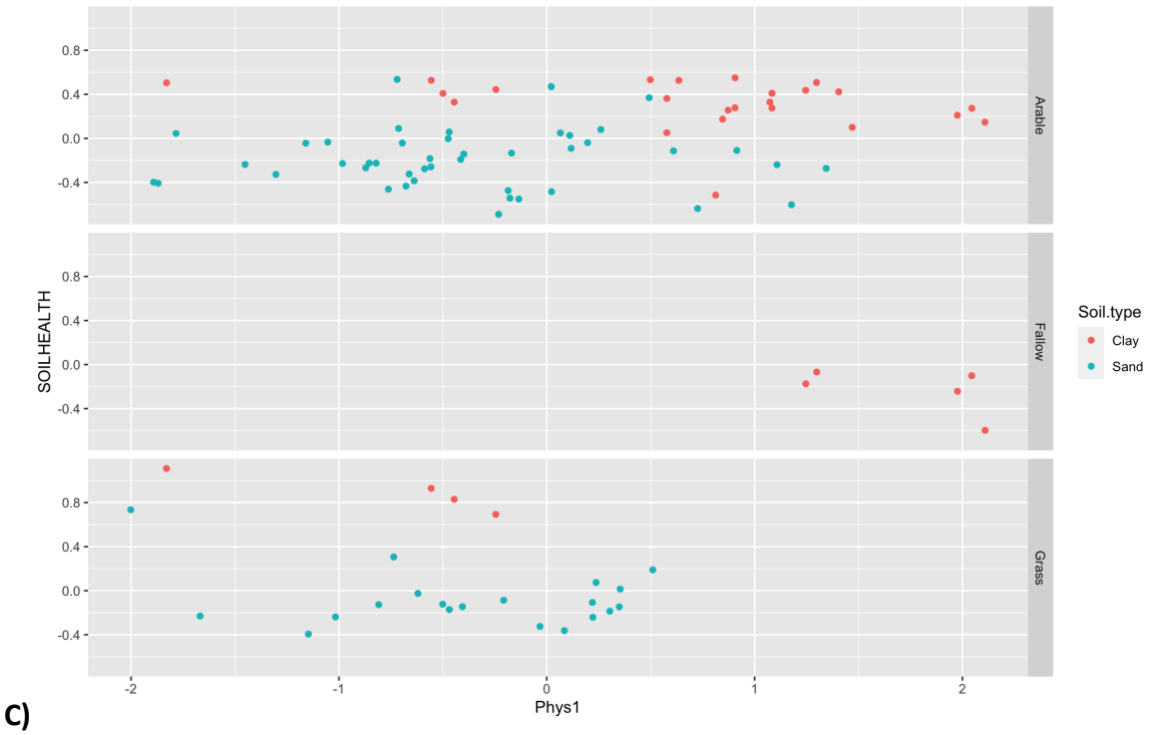
In Figure 20, we see that chemical soil parameters under arable, grass, and soils with no added amendments had differing effects on soil health of -0.2*, 0.73, and -0.15*, respectively. These differences confirm our hypothesis and highlight that a lack of second-order interactions may be giving a poor fit in our full model. Thus, soil type and land use history significantly impacted our measured soil health indicators, as evidenced by the regression results in Figures 19 and 20. To further develop our model, the next step would be to add more data to account for the interaction steps and effects. As we were limited to the sample size in our current dataset, we propose further data collection to make the proposed model more robust.



A)



B)



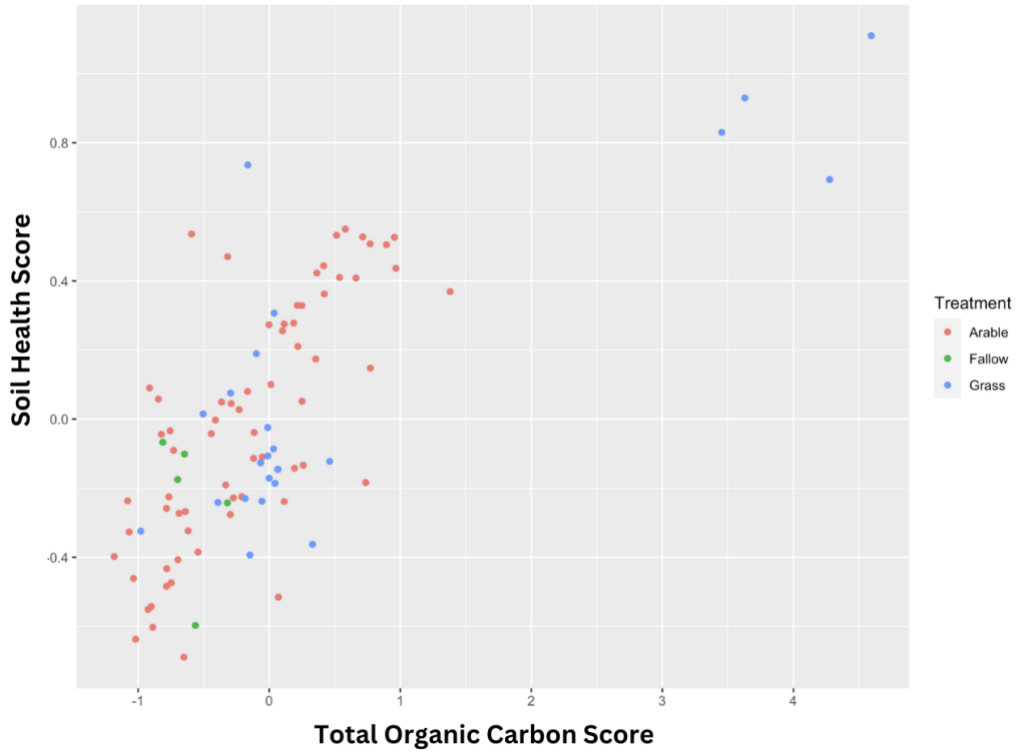
C)

Figure 21: Relationship between the soil health score and the A) Biological, B) Chemical, and C) Physical latent variable scores generated from our model across different soil types and land management practices.

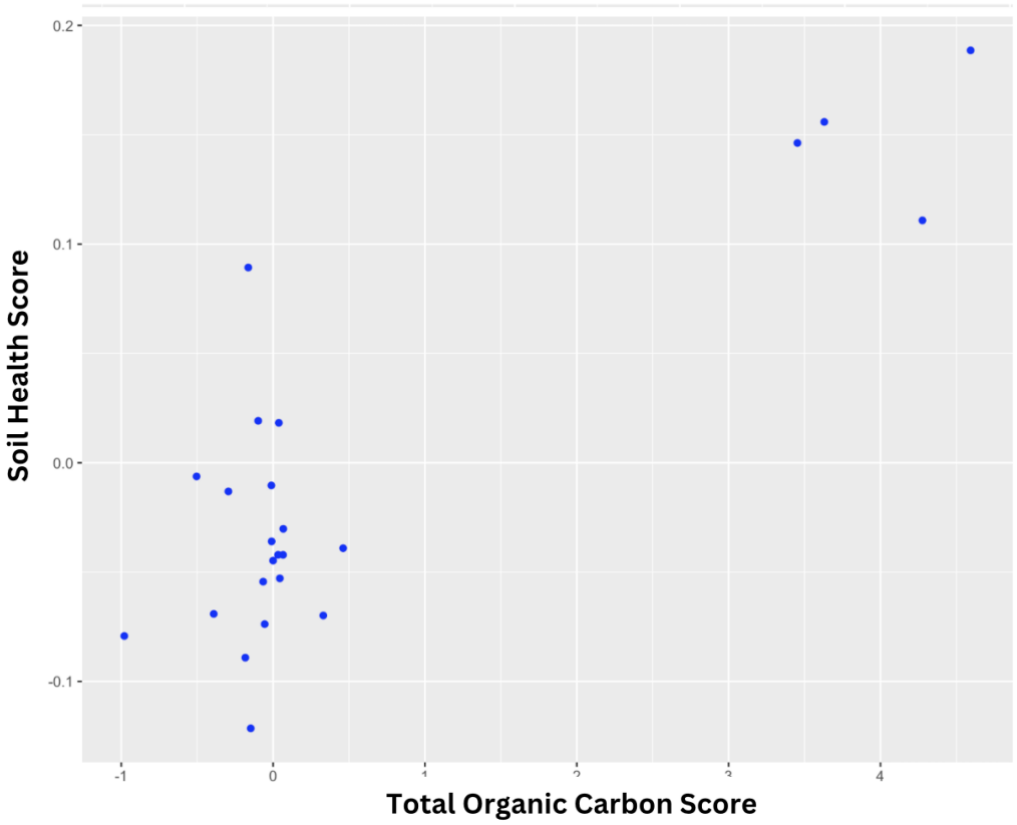
Figure 21 shows the relationships between our soil health score and the biological, physical, and chemical latent variables. The values of these latent variables were predicted from our full model (Figure 18). From Figure 21, we can see smaller soil health scores under fallow treatments, followed by grass and arable treatments. Furthermore, there were noticeable differences in the physical scores between grass and fallow treatments, indicating distinct structural and physical differences in soils under grass compared to fallow. However, it was hard to discern differences between grass and arable treatments based on chemical, physical, and biological scores.

We also observed a stronger negative relationship between our soil health score and biological score under grass than we do for arable. However, this was not seen across our physical and chemical scores, providing evidence for the need for additional data to allow our model to incorporate interactions based on land use and soil type. Furthermore, we could see clear differences in soil health scores for clay and sandy soils across biological, chemical, and physical variables, with soil health scores being greater for clay soils compared to sandy soils.

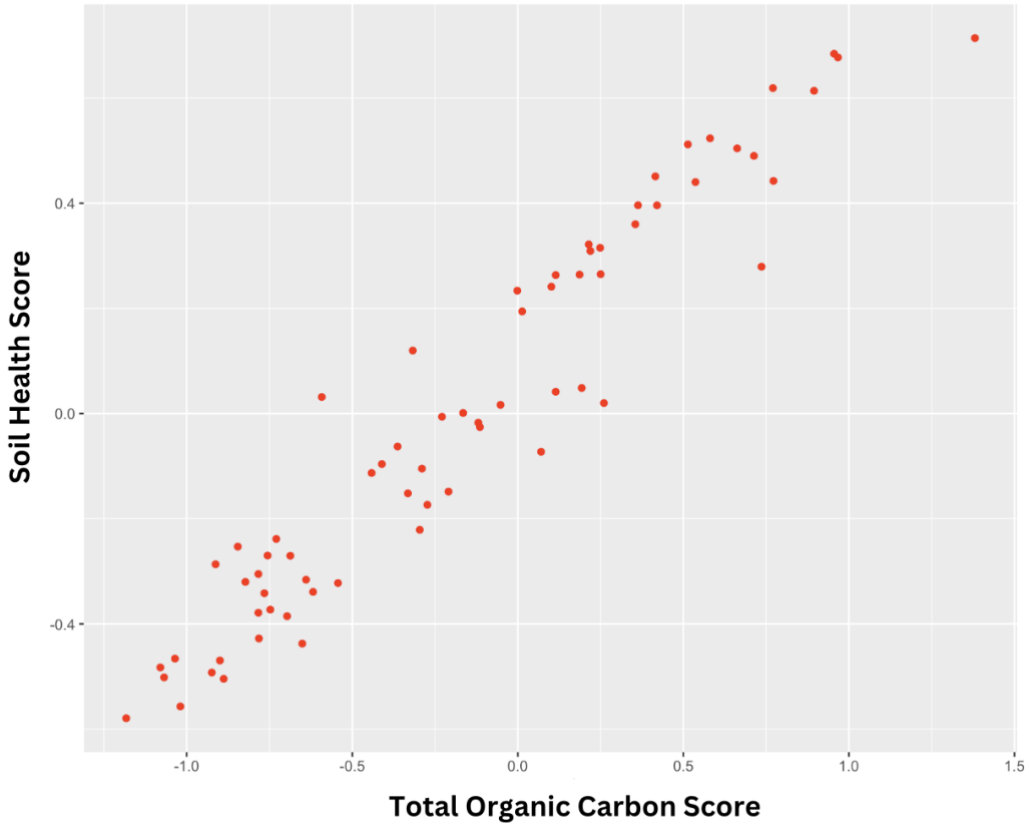
Overall, the relationship between both physical and biological latent variables and soil health was weak, whilst the chemical variable was negatively correlated with soil health (Figure 18, regression coefficient values). When we separated the predicted scores by land use and soil type however, we could see that the weak overall relationships may be due to opposing results across these landscape and management factors.


Soil Health Score
0.8
0.4
0.0
-0.4
-1
0
1
2
3
4
Total Organic Carbon Score
Treatment
Arable
Fallow
Grass

A)


Soil Health Score
0.2
0.1
0.0
-0.1
-1
0
4
Total Organic Carbon Score

B)



C)

Figure 22 illustrates A) Soil health scores derived from our structural equation model plotted against total organic carbon for all observed data points across grass, arable, and fallow treatments. Additionally, soil health scores generated from our modified, exploratory SEM, adjusted using soil data exclusively from B) Grass and C) Arable treatments, are presented against total organic carbon.

Figure 22A displays our model-generated soil health scores against organic carbon for our SEM fitted to the full dataset of observed variables from all land uses and soil types, whilst Figures 22B and 22C show the same relationships built from our exploratory models fitted to our subsetted data from under grass and arable treatments, respectively. We observed a stronger correlation between soil health scores under grass and arable treatments in this subset compared to soil health scores under the whole dataset, which was what we anticipated based on the results from Figure 20. However, although we observed positive correlations between total organic carbon scores and SI scores under grass treatments as well as total organic carbon and SI scores under arable treatments, a notable difference can be seen in the strength of these correlations. Specifically, arable treatments correlate more

strongly with the soil health score than grass treatments. Differences in correlations between grass and arable datasets underscore the importance of segregating our current model by land use and soil type and emphasise the need for additional data points to enhance model fitting and accommodate interaction steps and effects.

4.5 Discussion

Researchers have recognised the importance of incorporating a variety of biological, chemical, and physical properties in soil health assessments (Karlen et al., 2003). While there is widespread agreement on integrating these indicators into a comprehensive approach to understanding soil health, there is often ambiguity regarding their association with specific processes (Lehmann et al., 2020). Moreover, determining the relative importance of each indicator to the overall score is not always straightforward, although some researchers advocate for assigning equal weight to each sector (biological, chemical, or physical) (Cherubin et al., 2016). Using SEM, we were able to refine our minimum dataset of soil health indicators and categorise them into the appropriate groupings. SEM offers an advantage over standard PCA by providing deeper insights into how we conceptualise and categorise soil health indicators. In our study, we used SEM to define latent variables (Table 8) to describe soil health based on a minimum dataset of soil health indicators (Table 7). SEM then determined the weight of each soil health indicator based on the strength of relationships associated with total organic C, total N, and yield data, allowing for a more data-based approach.

Figure 18 presents our SEM for soil health, integrating environmental, management, chemical, physical, and biological factors. Environmental variables such as soil moisture, temperature, and soil type (sand content) significantly impact soil health through their effects on chemical, biological, and physical components. Soil moisture and temperature are key environmental variables that influence microbial activity, organic matter decomposition, and nutrient cycling. For example, soil moisture regulates microbial respiration, while temperature affects enzymatic activities involved in nutrient cycling (Schimel et al., 2007). Sand content influences soil texture, water retention, and aeration, directly affecting soil biological activity and nutrient availability (Six et al., 2004).

Moreover, management practices, including organic amendments, grass, and arable land use, affect soil health primarily through biological and chemical mechanisms. Organic amendments can increase SOC and stimulate microbial activity, while grass cover can enhance soil structure and prevent erosion. In contrast, arable management practices often lead to soil compaction and nutrient depletion, reducing biological activity and degrading chemical properties (Lal, 2004). SEM modelling in Figure 18 shows direct effects on biological factors, confirming the importance of management in controlling microbial activity and nutrient availability.

Environmental and management factors influence soil health through their impact on chemical (Chem1, Chem2, Chem3), biological (Bio1, Bio2), and physical (Phys1) properties. The chemical properties of soil, such as nutrient content and pH, are crucial for determining soil fertility and plant growth. Microbial activity (Biological factors) drives nutrient cycling and organic matter decomposition, while physical properties like bulk density influence root growth and water retention (Gregorich et al., 2000). SEM analysis suggests that both management and environmental factors act on these intermediate variables before ultimately affecting soil health.

Chemical (Chem1, Chem2, Chem3) and biological (Bio1, Bio2) factors have direct influences on overall soil health. Chemical properties such as nutrient levels and pH are directly related to soil fertility, which in turn impacts crop yield and organic matter turnover. The biological health of the soil, indicated by microbial biomass and enzyme activities, directly contributes to nutrient cycling and organic matter decomposition (Dick, 1994). In our model, the biological components are shown to have stronger direct effects on soil health than chemical factors, highlighting the role of microbial activity.

Furthermore, our SEM revealed that physical properties (Phys1) played a moderating role in the relationship between soil health and its drivers, showing a relatively weak direct influence. While soil physical properties like bulk density, soil texture, and aggregation are important, they often exert indirect influences on soil health. They regulate water retention and root penetration, which in turn affect biological and chemical processes (Tisdall et al.,

1982). Our model suggests that physical properties do not directly drive soil health but moderate other relationships, such as the impact of chemical and biological factors.

Soil health is ultimately represented by key indicators such as organic carbon, yield data, and total nitrogen. SOC is widely recognised as a primary indicator of soil health due to its role in supporting microbial activity, nutrient retention, and soil structure (Lal, 2004). Similarly, nitrogen availability is a critical nutrient that determines plant productivity and overall soil fertility. Crop yield is an integrative measure of soil health, as it reflects both the physical and biochemical conditions of the soil (Gregorich et al., 2000). The strong correlations shown in the model reflect the importance of these variables in determining soil health outcomes. We have found that our SEM builds on established relationships between environmental, management, chemical, biological, and physical factors, whilst revealing that soil health is primarily driven by biological and chemical factors, with management practices and environmental variables playing indirect roles. Thus, SEM has facilitated our understanding of how these factors interact, providing a robust framework for assessing soil health across different contexts.

The conventional additive approach for soil health scoring involves assigning individual scores to different soil health indicators and summing them to obtain an overall soil health score (Rinot et al., 2019). This method typically treats each indicator as equally important and assumes that their effects on soil health are independent of each other. It does not account for potential interactions or correlations between indicators (Schulte et al., 2006). The SEM approach offers several advantages over conventional additive soil health scoring approaches. Unlike the additive approach, SEM allows for the incorporation of latent variables, which may better capture the underlying complexity of soil health dynamics. Additionally, SEM can assess the relationships between different soil health indicators, providing insights into how they interact to influence overall soil health (Karlen et al., 1994). Moreover, SEM offers a more rigorous statistical framework for analysing complex datasets, potentially yielding more robust and reliable results compared to the simplistic additive method.

Congreves et al. (2015) found that when detecting changes in soil health across different land management, PCA-weighted soil health parameters exhibited up to ten times greater sensitivity than the conventional additive function. Similarly, Mukherjee and Lal (2014) found strong correlations among additive functions, weighted additive functions, and weighted PCA-modelled scores. However, they favoured the weighted PCA-modelled scores due to their objectivity and relatively higher correlations with crop yield. Furthermore, Askari and Holden (2014) reported similar scores between additive and weighted additive approaches but noted that the weighted additive approach, based on a minimal indicator dataset, exhibited the highest discriminatory power among all indices (Maaz et al., 2023). Together, these results suggest an increasing consensus towards using PCA-weighted indicators over additive functions in soil health assessments.

To serve as an effective tool, a scoring function must be capable of identifying changes in management practices or land use. We found that our soil health score showed a stronger negative relationship between our soil health score and biological score under grass compared to arable (Figure 21). This could be explained by the increased levels of soil organic carbon (SOC) found under grass, which not only serves as an indicator of soil health but also facilitates increased microbial activity (Kallenbach et al., 2016). Furthermore, SOC tends to decrease during land use changes from natural ecosystems to agricultural systems, which would explain the smaller biological score under arable systems compared to grass (Omonode et al., 2006). However, we did not detect any correlations between our soil health scores and physical and chemical scores across different land management (Figure 21). Similarly, Saviozzi et al. (2001) found that intensive cultivation resulted in a decline in all measured soil health indicators. As a result, they found grasslands showed greater soil health scores compared to arable sites. Furthermore, according to Noponen et al. (2013), disparities attributed to differences in land use or management may vary with depth. This is potentially significant given that most of our soil samples were collected up to a depth of 15 or 25 cm, depending on the observed variable. We suggest that further data would allow our model to better incorporate interactions based on land use.

One of our main objectives was to devise a scoring mechanism capable of quantifying the impacts of land use changes and management practices on soil health. However, our

proposed SEM designed for the soil health index showed a significant chi-squared value, suggesting discrepancies between the proposed model and the observed data. Several potential issues could be causing this discrepancy. For example, there may be errors or inconsistencies in the observed data used to fit the model. The proposed model, therefore, would not accurately represent the underlying relationships among the variables in the dataset, leading to poor model fit. During our analysis, we noticed a gap in our current database in the representation of soils collected from fallow and grass systems, as well as organic amendment plots. Thus, we were not able to fully examine the effect of management within these land uses. Furthermore, a major limitation of this study was that the assessment of latent variables (Table 7) hinges on the data we gathered and, therefore, must be understood within that framework. However, we found that most of our measured soil health indicators showed significant results across soil types and land uses but not across organic amendment additions (Table 9), leading us to believe that inconsistencies within our observed data were not solely responsible for the poor fit of our model. Moreover, we identified that one main issue was that our proposed model was too complex relative to the available data, resulting in poor generalisation of new data and unreliable parameter estimates.

We also discovered that the relationship between our model-generated soil health scores and organic carbon for our SEM fitted to the full dataset of observed variables from all land uses and soil types showed a weaker positive correlation compared to exploratory models fitted to our subsetted data from grass or arable treatments (Figure 22). We observed a stronger correlation between soil health scores under arable treatments in this subset compared to soil health scores under the whole dataset. These differences underscore the importance of segregating our model by land use and soil type and emphasise the need for additional data points to enhance model fitting and accommodate interaction steps and effects. We believe that addressing these issues by improving our data quality and increasing the sample size will improve our model fit to the data.

Maaz et al. (2023) also attempted to design a soil health scoring function using SEM. Surprisingly, their investigation revealed that soil health indicators did not conform to traditional biological, chemical, and physical categorisations within the SEM. They also found

that SEM was more effective in scoring soil health compared to conventional scoring methods, demonstrating the superior discriminatory power of SEM. Moreover, Maaz et al. (2023) generated a scoring function that was able to address intrinsic soil properties crucial for accurately assessing modern management practices across diverse soil types, encompassing tropical and volcanic soils and varying land use histories. Unlike our SEM, their model was able to highlight differences across management practices and soil types despite being fitted to the full dataset of observed variables.

A reliable soil health index must account for soil heterogeneity across time and space, as well as adhere to standardised soil sampling methodologies and analysis protocols. Additionally, the reliability of the index is influenced by model limitations associated with indicator selection and underlying assumptions used in its evaluation. Due to inherent variations and diversity among soils, standard practice for existing scoring indices defines soil health indices separately across various soil categories, often based on textural classification (Karlen et al., 2003). SEM can allow us to see the relationships between measured soil health indicators to determine the weight of each attribute or group of attributes in order to develop an integrated soil health score across different land uses and managements (Rinot et al., 2019). This study emphasises the potential of our proposed model and future soil health scores that utilise SEM in soil health assessments.

4.6 Conclusion

While many studies assess individual soil health indicators, only a limited number provide a comprehensive soil health score. In this study, we adopted a novel methodology utilising SEM to generate a soil health scoring function that can be used across a variety of soil types and land uses based on a minimum dataset of soil attributes. By employing SEM, we were able to capture the complex interactions among biological, chemical, and physical properties of soils. Whilst our model was not able to differentiate between different land uses and soil types when using our entire dataset, we did observe differences when using a subset of our data, showing that soil type and land use history significantly impacted our measured soil health indicators. Thus, our findings underscore the importance of considering the inherent variability among soils and the need for tailored approaches to soil health assessment based

on specific land use and soil types. Furthermore, our study highlights the limitations of traditional soil health indices and the necessity for a more integrated and comprehensive approach that accounts for the diverse range of soil functions and processes. Through the integration of various soil health indicators and the incorporation of latent variables, our model offers a more holistic understanding of soil health dynamics.

Moving forward, future research endeavours should focus on refining and validating our proposed soil health scoring function across different soil types and land management practices. Additionally, efforts should be directed towards elucidating the underlying mechanisms driving soil health processes and interactions, thus informing more targeted and effective soil management strategies. Overall, SEM provides a powerful analytical tool that aids in understanding the intricate relationships between different variables to help develop and test models that help unravel the complexity of soil systems, which will undoubtedly contribute to informed decision-making. This research thereby contributes to the ongoing development of soil health assessments and underscores the importance of adopting innovative approaches to better understand and manage our soils sustainably.

5 Main Discussion

Throughout this thesis, the significance of design choices when creating a soil health index was emphasised. This involves selecting the appropriate methods for measuring soil health indicators and exploring novel methodologies (Chapter 2), as well as identifying potential soil health indicators and validating their use in soil health assessments (Chapter 3). The overarching objective has been to leverage these insights to create a soil health index that is both robust and reliable (Chapter 4). The following discussion will delve into the implications and broader insights gathered from the primary findings.

5.1 Selecting the Best Method to Measure a Soil Health Indicator

Soil aggregate stability reflects the ability of soil aggregates to withstand disintegration when subjected to external forces such as rainfall, irrigation, or tillage (Amézketa 1999). Stable aggregates are crucial for maintaining crop productivity, reducing soil erosion, and promoting soil health (Neal et al., 2020). Thus, aggregate stability serves as a pivotal indicator for soil health and erosion risk assessment. However, its effective use in these assessments is impeded by the absence of convenient, standardised procedures, posing significant challenges in its use beyond research.

There are several traditional ways to measure aggregate stability. For the wet sieving technique, soil aggregates are submerged in water and gently agitated to separate them into different-size fractions based on their stability. The stability of aggregates is inferred from the amount of material retained on sieves of various mesh sizes after wet sieving (Kemper et al., 1986). Critiques of the wet sieving method highlight its limitation in capturing all mechanisms contributing to aggregate breakdown in natural field conditions (Kemper et al., 1986). In response, Le Bissonnais. (1996) proposed a more comprehensive approach known as the 'unified framework.' This method integrates the use of both water and ethanol as wetting fluids, along with varying rates of wetting (slow and fast) and mechanical energy (shaking after pre-wetting) (Almajmaie et al., 2017). These conventional methods for measuring aggregate stability often require significant time, expense, and laboratory resources.

In Chapter 2, we explored a novel technique to measure aggregate stability called “The Soil Aggregate Stability (SLAKES) smartphone application (Farjardo et al., 2016).” It aims to quantify aggregate stability by measuring how quickly soil aggregates disintegrate once submerged in water. With just a Petri dish, three soil aggregates ranging from 2 to 15 mm, and water, along with the SLAKES application installed on a smartphone, users can obtain aggregate stability measurements within 10 minutes. Thus, SLAKES represents a convenient, rapid, and cost-effective approach to assessing aggregate stability compared to traditional methods. Notably, SLAKES eliminates the need for laboratory facilities and is accessible to anyone with a smartphone.

To evaluate the tool, we collected SLAKES measurements across a range of land uses and soil types and compared results with the Le Bissonnais method. The results showed that the Le Bissonnais method effectively distinguished between almost all treatments, where soil aggregates in grass-treated plots exhibited significantly higher stability compared to those in arable and fallow treatments under clay and sandy soils (Figure 2). While the SLAKES app accurately differentiated the highly stable aggregates in the grass treatment from the less stable ones in arable and fallow treatments, it did not achieve the same level of statistical significance between arable and fallow treatments as the Le Bissonnais method and was significantly less sensitive under sandier soils (Figure 3). Thus, the study showed that the SLAKES method was able to determine aggregate stability to some extent, but it was less sensitive than the traditional Le Bissonnais method. However, depending on the purpose, the SLAKES method can provide some measure of aggregate stability quickly and at low costs.

Flynn et al. (2020) examined the SLAKES application's ability to detect variations among conventional tillage, no-tillage, and perennial grass management practices, contrasting its performance with the Cornell wet aggregate stability test (CWAST). Their findings revealed SLAKES' superior sensitivity in distinguishing tillage management practices compared to CWAST. In contrast, our laboratory-based method showed heightened sensitivity in detecting differences between land management practices relative to SLAKES. Similarly, Rieke et al. (2022) conducted a comparative study of the SLAKES app with three other aggregate

stability tests. They found SLAKES had a weaker correlation with these tests and was limited in its ability to discern differences in aggregate stability across treatments. However, our investigation demonstrated a significant correlation between SLAKES and the Le Bissonnais method.

Each methodology for assessing aggregate stability applies different types and levels of disruptive energy, leading to variations in how closely they reflect the disruptive forces encountered by soil aggregates in their natural environment (Almajmaie et al., 2017). Consequently, establishing a standardised method for measuring aggregate stability faces many problems, as relying solely on one approach may not adequately capture the behaviour of aggregates in real-world settings. Thus, selecting a methodology should prioritise replicating the types and magnitudes of disruptive energy experienced under natural field conditions. For example, aggregates subjected to rapid immersion, such as those encountered during flood or irrigation events, would be best evaluated using the wet sieving method (Truman et al., 1990). Ideally, all analyses should be performed using aggregates at moisture levels comparable to those observed before irrigation or rainfall events. Alternatively, analyses can be conducted using aggregates at air-dried moisture content levels as a standard practice.

As shown in our study, the cheaper and easier approaches to measuring aggregate stability are not the most sensitive at determining treatment differences across soil types compared to more conventional techniques. So, while each method has its own advantages and limitations, when selecting the most appropriate way to measure aggregate stability, we should consider a number of factors, such as soil characteristics, the nature of disruptive forces experienced by soil aggregates, and the availability of resources and equipment, before making a decision. Ideally, the chosen method should be simple, cost-effective, reproducible, and capable of distinguishing between different soils and treatments.

5.2 Selecting Potential Soil Health Indicators in Soil Health Assessments

Soil enzymes (EE) are crucial in ecosystem functioning. Researchers have proposed using enzymes as soil health indicators in agroecosystems due to their vital role in delivering

ecosystem services and soil functioning (Karaca et al., 2010). EE are involved in the decomposition and transformation of OM, releasing plant-available nutrients, and participating in C, N, and P cycling. Despite this, enzyme activity was found to be one of the least common soil properties used in soil assessments compared with other biochemical parameters (Saviozzi et al., 2002). However, in the last decade, soil EEs have been increasingly used as soil health indicators, allowing land managers to understand and monitor ecological changes in their soils.

Chapter 3 aimed to investigate how N-acetyl-β-glucosaminidase (NAG), acid phosphatase (PHO) and β-glucosidase (GLU) respond to different land uses and relate with other soil properties and evaluate the use of EE as a helpful indicator of soil health. The study also examined the impact of soil organic amendment (OA), soil type, and land management practices on enzyme activities. The findings indicated that there were significant relationships among the activities of the chosen soil enzymes: GLU, NAG, and PHO (Fließbach et al., 2007). It was also found that soil OA, soil type, and land management practices had an impact on soil enzyme activities. Furthermore, the results supported the hypothesis that soils with increased SOC have increased enzyme activity.

Dale et al. (2001) suggested parameters for potential soil health indicators. They must be: 1) easily measurable; 2) sensitive and responsive to environmental stressors, natural disturbances, and changes over time; 3) able to predict changes in soil ecosystems; and 4) variable in response. This study showed that overall, PHO, NAG, and GLU activities were easy to measure, sensitive to land use changes and OA additions, and able to predict other essential soil chemical health indicators. The results confirmed that land uses with the least amount of soil disturbances have greater soil health, as grass treatments relative to all other plots showed increased levels of enzymatic activity, followed by arable and fallow, respectively (Figures 6 to 8). Also, soil EE activity correlated with other observed soil health indicators. In particular, NAG and PHO were correlated to more chemical attributes of the soil (pH, inorganic-P, total N and soil organic carbon (SOC) (Figure 15). However, GLU was the most sensitive to land management and OA additions (Figure 11).

Further on, the study highlights the significance of studying long-term experiments with known histories and large datasets in order to better comprehend the drivers of soil health.

We concluded that EEs are promising indicators of soil health and are important in maintaining or increasing SOC in agricultural lands. These findings also suggest that EE activity could serve as an early indicator of soil degradation, allowing land managers to implement management strategies for enhancing soil health in agricultural systems. However, it must be stressed that EE activity should not be used as a standalone measurement to describe soil function and health. Results should be compared against baseline values or reference conditions for meaningful assessments of soil health. Additionally, using a combination of indicators, including physical and chemical properties, alongside enzymatic assessments would provide a more comprehensive understanding of soil health.

EE assays, while valuable tools for studying microbial activity and nutrient cycling in soils, have some limitations. Soil is disturbed during the collection, preparation, and storage of samples, which can change enzyme activity. Measuring enzymes within 24 hours after sampling is optimal; however, this is only sometimes plausible, and samples must be stored. According to Lee et al. (2007), keeping fresh samples cool at 5 °C instead of freezing during transport and storage has less impact on enzyme activity for temperate soils. For long-term storage, freezing samples at -80 °C is better than drying for organic soils (Wallenius et al., 2010). Storing soils after air-drying reduces enzyme activity and should be avoided (DeForest, 2009). Researchers agree that no clear standards can be made regarding the optimal storage of soil samples for enzyme tests because temperature and the duration of storage are enzyme- and soil-specific (Burns et al., 2013). As a result, soils from different climates and ecosystems would have different storage requirements. For example, Turner et al. (2010) recommended that temperate soils be stored at 5 °C, but tropical soils be stored at 22 °C. Furthermore, the time of year when soil samples are taken should be standardised. For example, attempts should be made to sample mid-season rather than just after the application of amendments or fertilisers to be more representative.

An unexpected problem encountered during the data analysis was that the protocols for measuring EE activity are not universal, making the comparison of studies difficult. There are two main approaches towards the enzyme assay methodology, which have the potential to produce remarkably different results (Dunn et al., 2013). The classical approach determines

enzymatic activity using substrate dilutions and optimises conditions such as pH and temperature. The second '*in situ'* approach attempts to replicate the natural conditions of the soil to estimate actual enzymatic activity (Burns et al., 2013). The two approaches are not mutually exclusive, and the choice of method is dependent on the research in question (Wallenstein et al., 2002).

An inherent difficulty of studying EE is that only a fraction of the total enzymes found in the soil can be assessed (Dodor, 2002). In our study, enzyme assays were performed under a strict set of conditions. Thus, the methods were clearly defined, and any changes to these conditions would change the measured activity. This means the enzyme assays measured the potential activity under optimal conditions and not the *in-situ* activity. Hence, the assays' conditions differ from those in the original soil, particularly the substrate concentrations that saturate the system.

Furthermore, EE can originate from plants, animals, and microorganisms, although it is widely considered that bacteria contribute the most to EE abundance. Enzyme assays cannot identify the source of enzymes in the soil. Instead, they measure the total activity of all enzymes present despite not all isoenzymes being assayed under their optimal conditions. As a result, enzyme assays may produce lower enzymatic activity readings (Knight et al., 2002; Nakayama et al., 2023). Despite these limitations, EE assays remain valuable tools for gaining insights into microbial processes and nutrient cycling in soils. Using a combination of assays and complementary techniques to address these limitations would result in a more comprehensive understanding of soil microbial activities.

Collaborations with other disciplines and technological advancements have already seen much progress in EE research. Enzyme assay methodologies are relatively quick and straightforward; as a result, they can be done on a routine basis. However, previous research recognises issues regarding the use of EE. Differences in methodologies and the lack of universal standards can complicate the interpretation of results and do not allow for comparisons across studies. Nonetheless, EE activities provide an integrative insight into changes resulting from external factors such as land management and the environment. Moreover, our ability to assess soil health and identify fundamental soil properties that can

serve as indicators of soil health has become a growing concern for land managers and farmers worldwide; the use of EE may be the solution to address these issues.

5.3 Integration of Measured Soil Health Indicators to Construct a Comprehensive Soil Health Index.

After selecting relevant soil health indicators and choosing which methodology is best suited to measure these variables, the next step is to weigh the contribution of each soil health indicator to quantify soil health. Using measurements of key soil health indicators collected from Rothamsted Research's long-term experiments across the UK, Chapter 4 aimed to design a soil health index using Structural Equation Modelling (SEM) capable of quantifying soil health across various land management practices and soil types. The results revealed that SEM offers a broad understanding of both observed indicators and latent variables contributing to overall soil health. However, the proposed model showed limitations in its applicability across diverse soil types and land management practices. Instead, its effectiveness was more pronounced when applied to specific land uses and soil types. Nonetheless, larger sample sizes are needed for a more comprehensive assessment. Despite these challenges, the study suggests that soil scientists could use SEM to refine soil health assessment models, enhance measurement accuracy, and deepen insights into the impacts of agricultural management practices on soil health.

Integrating measured soil attributes and their respective contributions into a soil health index posed a challenge during the initial design phase of our model. Current soil heath assessments commonly assign equal weight to each selected indicator, also known as an additive method (Fine et al., 2017). This approach may be oversimplified and may not reflect the complex contributions of different indicators to the soil system and soil functions. Gradual relative contribution, also known as a weighted additive, can be established through either a review of existing literature or expert opinions. The determination of the relative significance of each parameter is guided by specific objectives, such as productivity (yield or outcome), pesticide use, water use efficiency, and other relevant factors (Andrews et al., 2002; Wienhold et al., 2004). In addition, attributes may be further categorised into several soil functions, for example, water holding capacity, plant available water, soil degradation

resistance and nutrient supply power (Mukherjee et al., 2014). The relative importance of each soil function can be assigned either to equal or varying weightings. The differential relative weights could be based on the number of attributes within each soil function or determined by the perceived importance of each soil function. Greater relative importance might be attributed to field measurements, as they offer more pertinent and realistic data about the current state of the soil (Rinot et al., 2019).

An alternative approach is to categorise the measured attributes into sectors, such as physical, biological, and chemical, and then assign an equal weight to each sector. Kang et al. (2005) integrated soil attributes from three different indices (microbial activity, nutrient status, and crop yields) to quantify a unified soil sustainability index (Rinot et al., 2019). Nevertheless, the connection between particular soil attributes or soil categories and specific soil functions is subjective and currently lacks quantitative tools for validation. Moreover, concentrating solely on specific soil functions may prove inadequate for a system of such complexity. Hence, a more comprehensive quantitative approach is needed.

Another approach for integrating measured soil attributes and their respective contributions into a soil health index involves statistical analyses that outline the relative variability described by each parameter; this can include methods like principal component analysis (PCA) (Yu et al., 2018). Attributes incorporated into multidimensional scaling (MDS) for each soil sample can be assigned weights, based on the results from the PCA. Each principal component reveals specific proportions of variation found within the whole dataset. However, using this method could allocate a higher contribution to the most sensitive attribute, which may not necessarily be the same as the most important component in relation to soil functioning.

This study used PCA when creating our soil health index (Figure 17 and Appendix B), because PCA was able to condense multiple correlated soil health indicators into a smaller set of uncorrelated variables (principal components), which reduced the complexity of the model and prevented multicollinearity issues, which are often encountered with the additive weighting approach (Dormann et al., 2013). Furthermore, PCA objectively assigned weights to each principal component based on the variance explained by each component. In

contrast, additive weighting requires subjective decisions on the importance of each individual indicator, which can introduce bias and inconsistency across different studies. Finally, we found PCA was less sensitive to outliers and measurement errors and provided a clear interpretation of the relationships between soil health indicators and their contributions to overall variability, enhancing the soil health index's interpretability and was more robust in capturing the underlying structure of the data.

5.4 Challenges in the Development and Application of Soil Health Indices

There are several challenges and limitations in the development and application of soil health indices. Bünemann et al. (2018) conducted a review of major national soil assessment approaches that selected relevant indicators based on their conceptual relationships with soil functions and ecosystem services. They found that the most used indicators in soil health tests were pH, available phosphorous, and organic matter but noted that biological indicators were generally underrepresented. Moreover, Roper et al. (2017) examined common soil health tests in North Carolina and found that they were unable to differentiate between management systems or correlate soil health scores with measured yields. They concluded that focusing solely on specific soil threats, functions, or ecosystem services may not provide a comprehensive assessment of soil health. Furthermore, Roper et al. (2017) found that many indices primarily focused on soil productivity over limited growing periods, which may not accurately reflect the long-term sustainability of the soil (Rinot et al., 2019). However, van Es et al. (2019) reassessed their work and discovered the observed soil health indicators were correlated to crop yields and sensitive to management when measured by the Comprehensive Assessment of Soil Health (CASH). Their work highlights the inconsistencies in scoring approaches for quantifying soil health, which can result in the soil health index being susceptible to manipulation based on expert opinion or inappropriate statistical techniques, and further emphasises the need for standardised soil health assessments.

In this thesis, we have consistently observed significant differences across land uses and soil types, particularly between sandy loam and clay loam soils. These differences underscore the influence of soil type on soil health, indicating that soil health scores can vary not only

across different soil types but also within the same soil series under varying management practices. Nonetheless, we argue that a reliable index should encompass the diversity among soil types, land uses, and climate regions to effectively tackle the global challenges associated with soil health and sustainable agriculture. However, as we have discovered, accomplishing this goal is very challenging and would necessitate a significantly larger dataset and further research.

5.5 Future Works

While certain questions regarding soil EEs have persisted over time, fresh perspectives and innovations are now emerging to address them. Furthermore, the integration of novel technologies, such as the SLAKES app, could potentially revolutionise how we measure soil health and is anticipated to yield significant advancements. Ultimately, it is crucial to conduct extensive, field-scale research over time to successfully integrate indicators of soil health into more comprehensive biogeochemical models. We conclude this section by highlighting key research areas that could accelerate our understanding of EE and aid in designing future soil health metrics and indices.

1. Can EE activities observed in laboratory-based assays accurately reflect those in the field, and is it feasible to create nanoprobes that can be inserted into the soil for *in situ*, real-time measurements?

While briefly discussing the limitations of EE assays, the need for accurate EE measures was highlighted. Most EE assays are conducted *in vitro*, outside of their natural environment. This lack of *in situ* dynamics means that the assays may not capture the complex interactions and regulatory mechanisms that occur in the native soil environment. For example, enzyme activity is highly dependent on temperature and pH. EE assays are typically conducted under specific laboratory conditions, and the results may not accurately reflect the enzyme activities occurring in the field, where temperature and pH can vary widely. Therefore, we emphasise the need for *in situ*, real-time measurements that will better represent EE activities found in the field.

2. What is the potential impact of fluctuating environmental conditions on EE activity, and how will this affect ecosystem services?

EE activities are controlled by abiotic factors, such as temperature and pH, as well as biotic factors, including microbial activity and enzyme secretion. Global warming and shifts in precipitation patterns will likely cause drastic changes to ecosystem functioning, such as OM decomposition and nutrient cycling, and EE activity is likely to be receptive to these changes. The potential impact of fluctuating environmental conditions on EE activity raises concerns about its repercussions on ecosystem services. Variations in environmental factors can influence the ability of organisms to engage in EE, with subsequent effects on the delivery of essential ecosystem services. Understanding these dynamics is crucial for predicting and mitigating potential downturns in ecosystem services.

3. Is it possible to formulate a soil enzyme index that can also serve as an indicator of soil health?

A soil enzyme index would combine information from various soil enzymes to provide a comprehensive assessment of the soil's biological activity and functional capacity. Different enzymes play specific roles in nutrient cycling, OM decomposition, and overall soil ecosystem functioning. By measuring the activities of these enzymes, we can gain insights into the biological and biochemical processes occurring in the soil. Furthermore, a well-designed soil enzyme index would typically include enzymes associated with key soil functions such as C cycling (GLU), N cycling (NAG), and P cycling (PHO). The index may also consider ratios or combinations of enzyme activities to provide a more integrated view of soil health. A soil enzyme index would, therefore, allow for the integration of multiple soil processes, providing a more holistic view of soil health.

4. How can the methodologies for soil health indicators be effectively standardised to ensure accurate assessments of soil health?

As previously discussed, there are many contrasting methodologies that measure aggregate stability and EE activities, which ultimately influence our results and interpretations. Standardising the methodologies of soil health indicators, as well as sampling and handling

procedures in the field, is essential to ensure results can accurately reflect soil health status and predict soil functions. Without this, there is a risk of using misleading or inaccurate data, which could lead to incorrect assessments of soil health and ineffective management decisions. Moreover, standardised methodologies can ensure consistency and comparison of soil health assessments across different studies and locations. Addressing this can no doubt lead to further advancement of soil health research and the development of effective soil management strategies.

5. Can advancements in technology, such as remote sensing and molecular techniques, be leveraged to enhance soil health assessment and index development?

Advancements in technology hold significant promise for enhancing soil health assessment and index development. Remote sensing enables the collection of large-scale spatial data on soil properties and characteristics, facilitating comprehensive assessments of soil health over broad geographic areas. Molecular techniques, on the other hand, provide insights into soil microbial communities and their functions, offering valuable information on soil biological activity and health. By leveraging these technologies, researchers and practitioners can gain a deeper understanding of soil health dynamics and develop more robust soil health indices that account for various biological, physical, and chemical factors and ultimately allow for better-informed management decisions.

6. Can we develop a soil health index that is applicable across diverse soil types while accounting for variations in soil properties and management practices?

Much of Chapter 6 delves into answering this question by designing a soil health index that can be used across different soil types. However, further work is needed to refine and validate our soil health index to ensure its successful application across different soil types, land uses, and management practices. This involves conducting extensive field studies to collect more data from soil health indicators from diverse locations and ecosystems. Soil health varies significantly across different soil types due to differences in physical, chemical, and biological properties. Developing a soil health index that can effectively capture these variations and provide meaningful assessments across diverse soil types is crucial for

sustainable land management and agricultural practices, enabling researchers, policymakers, and land managers to make informed decisions and implement strategies to improve soil health.

6 Concluding Remarks

It is widely acknowledged that preserving soil health is crucial for human sustainability, and effective soil management is necessary to support human health and well-being while mitigating soil and environmental degradation. Thus, creating a quantitative, comparative test to evaluate the success of soil management approaches is a global challenge of high priority. An index that only caters to a specific soil type could fail to project soil health status on regional and global scales. Therefore, a reliable soil health index must be robust to assess different soil types despite the inherent variability in the soil.

Choosing the appropriate methodology for measuring soil health indicators is crucial for developing an effective soil health index. In Chapter 2, we evaluated a novel technique, SLAKES, alongside the Le Bissonnais method to assess aggregate stability. SLAKES showed sensitivity in differentiating management types on clayey soil, but its performance was limited on sandy soil. Nonetheless, SLAKES presents a promising option for quick and convenient assessment of aggregate stability. However, our results highlight the need for further refinement and validation of the SLAKES app before its widespread use in soil stability assessments.

Biological soil health indicators are often underrepresented in soil health assessments. Chapter 3 investigated extracellular enzymes (NAG, PHO, GLU) as potential soil health indicators. We found that grass treatments exhibited greater enzymatic activity compared to arable and fallow treatment, and enzyme activity correlated with other established soil health indicators, indicating their potential as comprehensive biological indicators of soil health. These results underscore the importance of considering biological processes in soil health assessments and highlight the need for further research into the role of enzymes in soil health dynamics.

Finally, Chapter 4 integrated insights from previous chapters to develop a meaningful soil health index using structural equation modelling (SEM). We found that SEM proved invaluable in comprehensively understanding complex soil health dynamics, particularly in interpreting relationships among multiple indicators and latent variables. Despite the challenges encountered in designing a soil health index that can be applied across diverse soil types, our research suggests that such an endeavour is feasible with further investigation and larger datasets. Ultimately, the development of a robust soil health assessment necessitates collaborative efforts and data collection from a wide range of soil types and management practices to ensure its effectiveness and reliability.

Throughout this thesis, the multifaceted nature of designing a soil health index was highlighted, emphasising the importance of selecting informative indicators, appropriate measurement methodologies, and suitable model design. By addressing these considerations, this research lays the groundwork for more effective soil health assessments.

Appendix A: Detailed Description of Long-Term Experimental Sites.

Highfield Ley-Arable Long-Term Experiment and Highfield Bare Fallow

Soil samples were taken from the Highfield Ley-Arable long-term experiment located at Rothamsted Farm in Harpenden, UK (geolocation: 51.802777, -0.366025). The experiment started in 1948 in a field that had been in permanent grass since 1838. In total, six treatments were established in a randomised complete block design with four replicate blocks. The treatments included permanent grass (i.e., continuation of the former land use), arable rotations, and some alternating between grass ley and arable rotations. In 1959, an area of permanent grass (since 1838) adjacent to the Highfield Ley-Arable experiment at Rothamsted was ploughed, and crops have not grown there since; this is the Highfield Bare Fallow. A similar long-term bare fallow (the 'Geescroft Soil Mine') is located adjacent to the Highfield Bare Fallow. These areas have been kept bare fallow by frequent cultivation, so inputs of carbon into the soil have been negligible.

In 2008, two plots (containing long-term continuous grass and arable treatments, respectively) within each block of the Highfield Ley-Arable experiment were split to establish (or maintain in the case of the original treatment) continuing grass, arable and bare-fallow treatments. Similarly, on the Highfield Bare Fallow area, new 'plots' of grass and arable treatments were established in 2008. Together, this split-plot establishment forms the Highfield Conversion experiment. We nevertheless focused on the split-plots continuing the long-term (since 1948 or 1959) grass, arable, and bare fallow treatments for this research. These treatments are all continuous with no rotation ('arable' is continuous winter wheat since 2008). For the bare fallow treatment, 5 'plots' were established in 2008 (3 on Highfield Bare Fallow and 2 on the Geescroft Soil Mine) to maintain the long-term treatment on what was formerly just single parcels of land. Therefore, in total, we focussed on 13 plots under continuing long-term treatment (8 under long-term arable or grass, 5 under long-term bare fallow) (Table A1 and Figures A1 and A2). For further information, visit the e-RA webpages: https://www.era.rothamsted.ac.uk/experiment/rrn1 and https://www.era.rothamsted.ac.uk/experiment/rrs1

Table A1: Current treatment components in the Highfield Ley-Arable long-term experiment and Highfield Bare Fallow used in our experimental study.

Treatment	Description of treatment
Rc(Rc)	Permanent grass/clover ley (since 1948)
A(A)	Arable rotation (1948-2008), continuing as permanent winter wheat (2008-)
F(F)	Permanent bare fallow (since 1959)

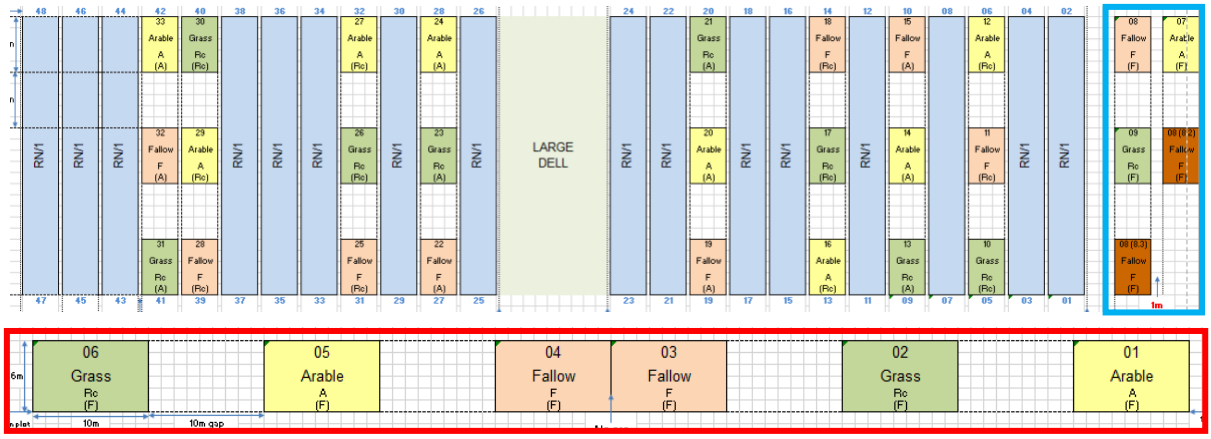


Figure A1: The current layout of the Highfield Conversion experiment showing split-plots on the former long-term grass and arable treatments of the surrounding Highfield Ley-Arable experiment and newly created plots on the former Highfield Bare Fallow (right of upper panel, outlined in blue), including the Geescroft Soil Mine (bottom panel, outlined in red) fallow areas.

Figure A2: Aerial view of the Highfield Ley-Arable experiment. The Highfield Bare Fallow is shown towards the top in relation to the Highfield Ley-Arable and Highfield Conversion (sometimes also known as 'Reversion') experiments. The Geescroft Soil Mine is above and to the right of the Highfield Bare Fallow, outlined in red.

Woburn Ley-Arable Experiment

This experiment, originally set up in 1938 at the Woburn Experimental Farm in Bedfordshire, UK (geolocation: 51.99906, -0.61673), tests the effects of continuous arable and ley-arable cropping on crop production, soil organic matter dynamics and fertility in a sandy loam soil. Since 2012, some changes were made, where there are essentially 4 five-year rotation treatments (2 arable rotations and 2 grass ley rotations), with 4 replications in the 16 plots (4 treatments x 4 replications). A five-year rotation comprises 3 years of 'treatment' (i.e., the arable or grass ley rotation) followed by 2 years of 'test' (under arable crops)
A main difference, therefore, with the Rothamsted Highfield Ley-Arable experiment is that there are no treatments which are permanently under grass (the 'grass ley' treatments all have two years under arable crops).

The two arable rotation treatments during the experimental period had recently been standardised (since 2021) and were in a 3-year treatment phase of annual winter crops of rye, barley, and oats (the order of these varied between the two treatments, as did one of the crops prior to 2021). The two grass ley rotation treatments are either 3 years under grass (with N fertiliser), or 3 years under grass-clover (with no N fertiliser). Following the 3 years under the treatment, there is the 2-year test phase, under annual winter crops of wheat and then rye. During the test phase, the 16 plots are subdivided into 4 subplots to test 4 levels of N fertiliser application (including a 0 rate) which rotate from cycle to cycle (i.e., after 4 cycles, each subplot will have received the same amount of N). The 16 plots were replicated in 5 blocks during the experiment (80 plots in total).

With 5 blocks and a 5-year rotation, it means that in any one year, each year of the rotation will be present (i.e., a block under the 1st year of the treatment, a block under the 2nd year of the treatment, a block under the 3rd year of the treatment, a block under the 1st year of the test, and a block under the 2nd year of the test).

On the plot plan for the site (Figure A3), any single plot will comprise a two-part code: the first part in parentheses tells you the overall rotation treatment, and the second part tells

you the crop that it is currently in (e.g. '(ABe)W' indicated it is the arable rotation (R-WB-O) currently under test year 1 (wheat).

For the purposes of soil health, we looked at three treatments, the newly uniform arable rotation (ABe/AO) and two grass ley rotations (Lc3/Ln3), both sampled at the end of their third year of treatment. This will be when soil organic matter is at its maximum in the grass rotation treatments before they go into 2 years of cropping. We also sampled plots in their first year of test crop that were previously under grass and arable treatments to capture two extremes of the rotation – after the maximum effect of the treatment and after the 're-setting' test crop years.

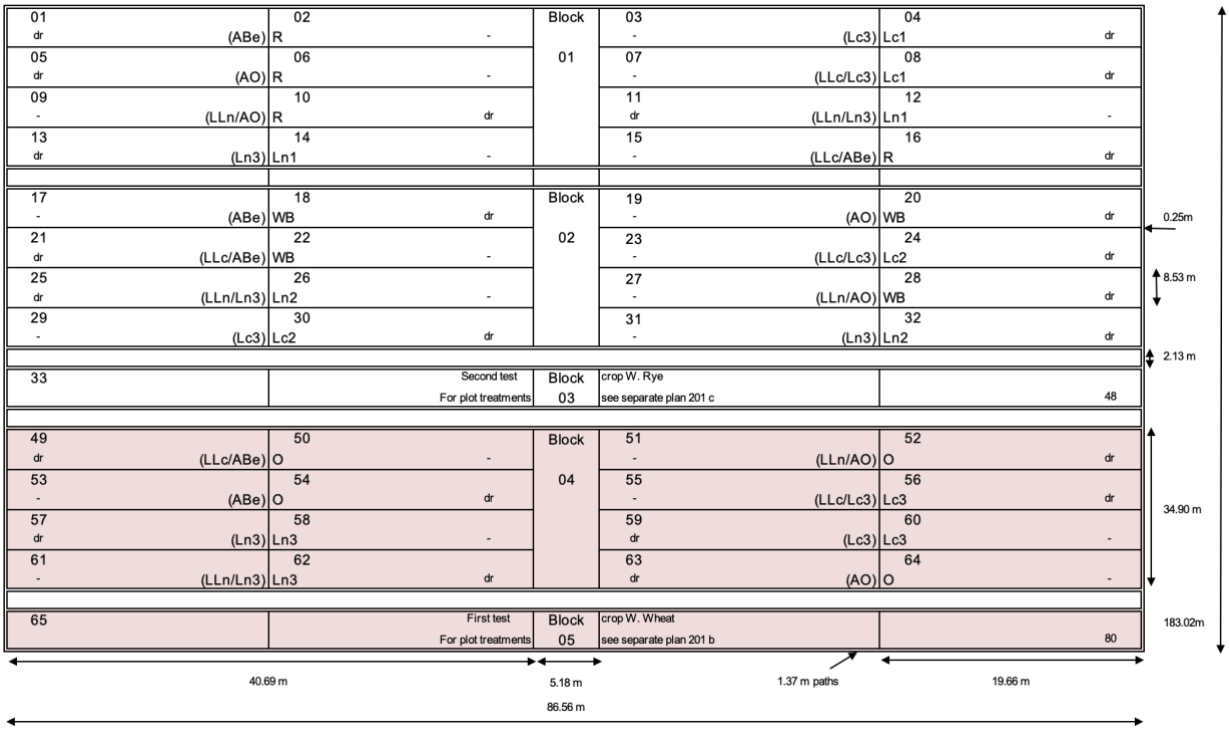


Figure A3: Plot plan for the site of the Woburn Ley-Arable experiment in 2022. Any single plot will comprise a two-part code: the first part in parentheses tells you the overall rotation treatment, and the second part tells you the crop that it is currently in. Blocks 4 and 5 (highlighted in red) were in their third year of treatment and first year of a test crop in 2022, respectively.

Table A2. The eight arable and grass ley rotations in the Woburn Ley-arable experiment. The rotation are five years, comprising three years under the arable or grass ley treatment, followed by two years under arable test crops. The code gives the overall rotation treatment in parentheses, followed by the current crop in that year. All arable crops are autumn-sown winter varieties. Note that changes were made in 2021 to align the arable rotations and to replace beans with barley.

Rotation treatment		Treatment year 1		Treatment year 2		Treatment year 3		Test year 1		Test year 2	
code	descriptor	code	descriptor	code	descriptor	code	descriptor	code	descriptor	code	descriptor
(ABe)	arable rotation	R	rye	WB	barley (formerly oats)	O	oats (formerly beans)	W	wheat	R	rye
(AO)	arable rotation	R	rye	WB	barley (formerly beans)	O	oats	W	wheat	R	rye
(LLc/ABe)	arable rotation	R	rye	WB	barley (formerly oats)	O	oats (formerly beans)	W	wheat	R	rye
(LLn/AO)	arable rotation	R	rye	WB	barley (formerly beans)	O	oats	W	wheat	R	rye

(Lc3)	grass ley rotation	Lc1	clover/grass	Lc2	clover/grass	Lc3	clover/grass	W	wheat	R	rye
(Ln3)	grass ley rotation	Ln1	grass	Ln2	grass	Ln3	grass	W	wheat	R	rye
(LLc/Lc3)	grass ley rotation	Lc1	clover/grass	Lc2	clover/grass	Lc3	clover/grass	W	wheat	R	rye
(LLn/Ln3)	grass ley rotation	Ln1	grass	Ln2	grass	Ln3	grass	W	wheat	R	rye

For further information, visit the e-RA webpage: https://www.era.rothamsted.ac.uk/experiment/wrn3

Woburn Organic Manuring Experiment

The Woburn Organic Manuring Experiment started in 1964 on a sandy loam soil at the Woburn Experimental Farm in Bedfordshire, UK (geolocation: 51.999805, -0.616036) to test the effects of different types of organic matter inputs on soil organic matter (SOM) and crop yields. The experiment has had three distinct phases of organic matter input, always with 8 treatments in a randomised complete block design with four replicate blocks (32 plots in all). Initially, six organic treatments (FYM, peat, straw, green manure and two grass leys) were compared with two mineral-fertilizer-only treatments. In 2003, the third and current treatment phase started, with 8 treatments (Table A3). An arable rotation (winter rye, spring barley, winter beans, winter wheat, forage maize) was started on seven treatments; the eighth treatment was sown to a grass/clover ley. The seven treatments under the arable rotation are split into 6 split plots to receive 6 levels of inorganic N inputs (except for the winter beans phase, which does not receive inorganic N), which rotates annually. In a seven-year period (accounting for the year of no N under winter beans), each split plot will have received the same amount of mineral N, and soil sampling conventionally is done across the whole main plot area to capture a representative sample across all split plots.

For further information, visit the e-RA webpage:
https://www.era.rothamsted.ac.uk/experiment/wrn12

Table A3: Current treatment components (since 2003) in the Woburn Organic Manuring experiment. Note that there are two F treatments – one is a long-term inorganically-fertilised treatment, the other has previously received organic amendment in a previous phase of the experiment.

Treatments since 2003	
Code	**Treatment**
F	None
Dg10	FYM at 10 t/ha annually
Dg25	FYM at 25 t/ha annually
St	Chopped wheat straw at 7.5 t/ha annually
CC	Cover crop (white mustard) prior to spring sown crop
Co	Compost at 40 t/ha annually
Lc	Permanent grass/clover at 30 kg/ha

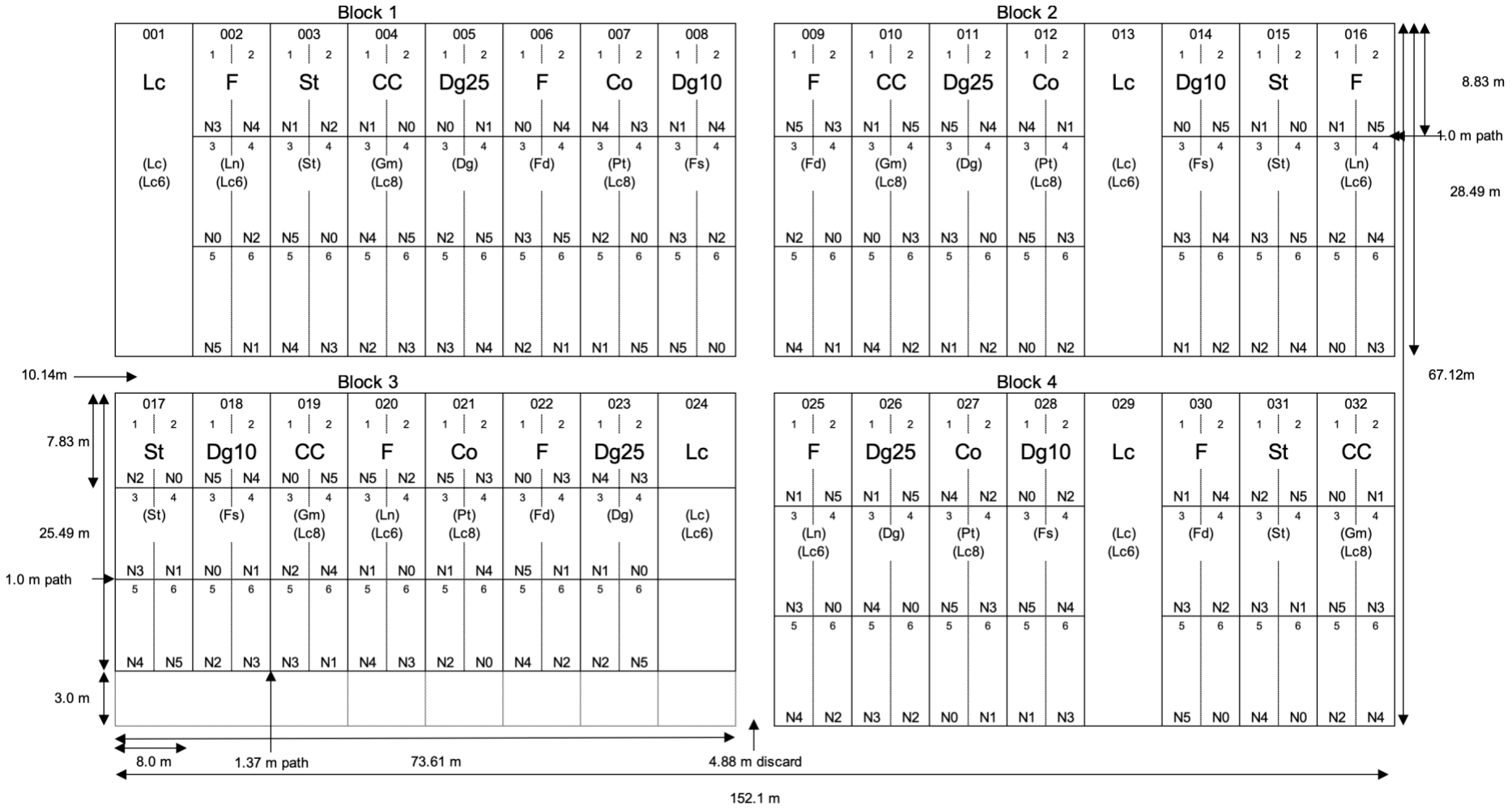


Figure A4: Plot plan for the site of the Woburn Organic Manuring experiment in 2022. Any single plot will comprise six different nitrogen additions: N0 to N6. All soil samples were taken from the N2 subplot from all plots.

Fosters Organic Amendment Experiment

On Fosters field at the main Harpenden site (geolocation: 51.812922, -0.379079), 220 ploughed plots tested five rates of addition of four kinds of organic matter (OM) amendment and 3 mixtures with straw and with the background N-response measured at 5 rates of N (2 crops x 2 blocks x (5 N rates x 4 OM types + 5 OM rates x (4 OM types +3 mixtures)) =220 experimental plots). Two arable rotation series were compared in two replicate blocks, with half the field sown with each crop in 220 plots (Table A4). The Fosters field experiment was located at Rothamsted Research farm (51.82 N and 0.37 W) in Harpenden, UK, which has a temperate climate in the South of England. The soil is characterised as a flinty clay loam of the Batcombe series, with a total organic C of 1.6 % and pH of 6.99. The trial is managed using a conventional regime (fertiliser, pesticides) and is tilled by ploughing. Both the organic amendments and nitrogen treatments (Table A5) were applied by hand each year in the autumn (farmyard manure was chopped first with a muck spreader).

Table A4: Cropping details on Fosters Organic Amendment field experiment.

Block	**2019**	**2020**	**2022**	**2023**
W Rotation 1	Winter Wheat, ww	Spring Barley, sb	Winter Oats, woats	Winter Wheat, ww
E Rotation 2	Spring Barley, sb	Winter Oil Seed Rape, osr	Winter Wheat, ww	Spring Barley, sb

Table A5: Treatments on Fosters Organic Amendment field experiment.

Organic matter	**Carbon rate (tonnes C ha^{-1})**	**Nitrogen rate (kg N ha^{-1}) N0-N4**
Straw	0, 1, 1.75, 2.5 or 3.5	N0-N4,
Anaerobic digestate		N0-N4
Anaerobic digestate + Straw		N3
Compost		N0-N4
Compost + Straw		N3
Farmyard manure		N0-N4
Farmyard manure + straw		N3

Appendix B: Chapter 4 Supplementary Information for Principal Component Analysis (PCA) Loadings.

Table B1: Standard deviation, proportion of variance, and cumulative proportion for Principal Component Analysis (PCA) related to chemical indicators.

	Principal component 1	**Principal component 2**	**Principal component 3**
Standard deviation	2.1661	1.5058	0.9860
Proportion of variance	0.5213	0.2519	0.1080
Cumulative proportion	0.5213	0.7733	0.8813

Table B2: Dimension loadings for chemical indicators, showing the contributions of each chemical indicator to the principal components.

Chemical indicator	**Dimension 1**	**Dimension 2**	**Dimension 3**
Aluminium	10.3193743	11.7750684	8.19237607
Magnesium	17.5242690	2.4334023	0.41206710
Manganese	17.9612971	0.2377030	5.03821049
Calcium	3.0006444	37.0207380	0.11864452
Iron	0.9696029	3.8951918	84.73454292
Potassium	11.8236801	14.1218771	0.03395387
Sodium	9.1885488	22.8806118	0.35299774
pH	13.5297683	0.1436069	0.10932442
Inorganic phosphate	15.6828149	7.4918006	1.00788287

Table B3: Standard deviation, proportion of variance, and cumulative proportion for Principal Component Analysis (PCA) related to biological indicators.

	Principal component 1	**Principal component 2**
Standard deviation	1.4684	0.9196
Proportion of variance	0.7187	0.2813
Cumulative proportion	0.7187	1.0000

Table B4: Dimension loadings for biological indicators, showing the contributions of each biological indicator to the principal components.

	Dimension 1	**Dimension 2**
β-glucosidase	46.24291	0.3440719
N-acetyl-β-glucosaminidase	19.55632	68.5392545
Acid phosphatase	34.20077	31.1166735

Table B5: Standard deviation, proportion of variance, and cumulative proportion for Principal Component Analysis (PCA) related to physical indicators.

	Principal component 1
Standard deviation	1. 1.6719
Proportion of variance	0.9317
Cumulative proportion	0.9317

Table B6: Dimension loadings for chemical indicators, showing the contributions of each physical indicator to the principal component.

	Dimension 1
Compaction	32.48751
Bulk Density	35.75055
Aggregate Stability	31.76194

Abbreviations

AD	Anaerobic Digestate
AMC	Amino-methyl coumarin
ANOVA	Analysis of Variance
BD	Bulk Density
C	Carbon
Ca	Calcium
CEC	Cation Exchange Capacity
CICES	Common Classification of Ecosystem Services
EE	Extracellular Enzymes
ES	Ecosystem Services
GLU	β-glucosidase
IQR	Interquartile Range
kPA	Kilopascals
L-DOPA	L-3,4-dihydroxyphenylalanine
LMM	Linear Mixed Model
Mn	Manganese
MWD	Mean Weight Diameter
MUF	Methylumbelliferone
N	Nitrogen
Na	Sodium
NAG	N-acetyl-β-glucosaminidase
OA	Organic Amendment
OM	Organic Matter
P	Phosphate
p	Probability
p-NP	p-nitrophenol
PCA	Principal Component Analysis
PLS	Partial Least Squares
PHO	Acid Phosphatase
SCD	Segmented Array Charged Coupled Detector

SI	Slaking Index
SH	Soil Health
SHI	Soil Health Indicators
SOC	Soil Organic Carbon
SOM	Soil Organic Matter
VIF	Variance Inflation Number
VNIR	Visible Near-Infrared Spectroscopy
VSA	Visual Soil Assessment

References

Abivena, S., Menasseri, S., Chenud, C. (2009) 'The effects of organic inputs over time on soil aggregate stability – A literature analysis'. *Soil Biology and Biochemistry*, 41. pp.1-12.

Adams, M.A., Buchmann, N., Sprent, J., Buckley, T.N., Turnbull, T.L. (2018) 'Crops, Nitrogen, Water: Are Legumes Friend, Foe, or Misunderstood Ally?'. *Trends in Plant Science,* 23(6). pp.539-550.

Adetunji, A.T., Lewu, F.B., Mulidzi, R., Ncube, B. (2017) 'The biological activities of β-glucosidase, phosphatase and urease as soil quality indicators: a review' *Journal of Soil Science and Plant Nutrition,* 17(3). pp.794-807.

AGS (2023) Responding to the UK's Soil Crisis - Sustainable Soil Management & the Future of Soil Reuse. Available at: https://www.ags.org.uk/2022/11/responding-to-the-uks-soil-crisis-sustainable-soil-management-the-future-of-soil-reuse/ (Accessed 15 December 2023).

Aldana-Jague, E., Heckrath, G., Macdonald, A., van Wesemael, B., Van Oost, K. (2016) 'UAS-based soil carbon mapping using VIS-NIR (480–1000 nm) multi-spectral imaging: Potential and limitations'. *Geoderma,* 275, pp.55-66.

Ali, J., Sharma, D.C., Bano, A., Gupta, A., Sharma, S., Baipai, P., Pathak, N. (2019) 'Chapter 29 - Exploiting Microbial Enzymes for Augmenting Crop Production', in Kuddus, M. (ed.) *Enzymes in Food Biotechnology.* London: Academic Press, pp.503-519.

Almajmaie, A., Hardie, M., Acuna, T., Birch, C. (2017) 'Evaluation of methods for determining soil aggregate stability'. *Soil and Tillage Research*, 167. pp.39-45.

Amacher, M.C., O'Neil, K.P., Perry, C.H. (2007) 'Soil vital signs: a newsoil quality index (SQI) for assessing forest soil health'. *US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, C*O. RMRS-P-65, pp. 65.

Anas, M., Liao, F., Verma, K.K., Sarwar, M.A., Mahmood, A., Chen, Z., Li, Q., Zeng, X., Liu, Y., Li, Y. (2020) 'Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency'. *Biological Research,* 53. pp.47-67.

Andrews, S.S., Carroll, C.R. (2001) 'Designing a soil quality assessment tool for sustainable agroecosystem management: Using the ecological process management approach'. *Ecological Economics*, 37(3). pp.355-369.

Andrews, S.S., Karlen, D.L., Cambardella, C.A. (2004) 'The Soil Management Assessment Framework'. *Soil Science Society of America Journal*, 68(6). pp.1945-1962.

Avery, B.W. (1980) *Soil Classification for England and Wales (Higher Categories). Technical Monograph 14.* Harpenden: Soil Survey of England and Wales, pp.67.

Avery, B.W., Catt, J.A. (1995). *The Soil at Rothamsted.* Harpenden: Lawes Agricultural Trust, pp.1-44.

Bandick, A.K., Dick, R.P. (1999) 'Field management effects on soil enzyme activities'. *Soil Biology and Biochemistry,* 31(11). pp.1471-1479.

Baveye, P.C., Baveye, J., Gowdy, J. (2016) 'Soil "Ecosystem" Services and Natural Capital: Critical Appraisal of Research on Uncertain Ground'. *Frontiers in Environmental Science,* 4. pp.1-49.

Beauchamp, E.G., Drury, C.F. (1991) 'Ammonium Fixation, Release, Nitrification, and Immobilization in High- and Low-Fixing Soils'. *Soil Science Society of America Journal,* 55(1). pp.125-129.

Beillouin, D., Cardinael, R., Berre, D., Boyer, A., Corbeels, M., Fallot, A., Feder, F., Demenois, J. (2022) 'A global overview of studies about land management, land-use change, and climate change effects on soil organic carbon'. *Global Change Biology,* 28. pp.1690–1702.

Berge, H.F.M., Schröder, J.J., Olesen, J.E., Cerva, J.V. G. (2019) 'Soil quality: a confusing beacon for sustainability'. *Proceedings International Fertiliser Society,* 841. pp.64.

Bibby, J.S., Mackney, D. (1969) *Land use capability classification*. Technical Monograph No.1. Harpenden: Soil Survey of England and Wales. pp.145.

Boltz, D.F., Mellon, M.G. (1948) 'Spectrophotometric determination of phosphorus as molydiphosphoric acid'. *Analytical Chemistry,* 20(20). pp.749-751.

Borrellia, P., Alewell, C., Alvarezde, P., Anache, J., et al. (2021) 'Soil erosion modelling: A global review and statistical analysis'. *Science of the Total Environment*, 780(1). pp.146494.

Brandt, J., Vejre, H. (eds.) (2003). *Multifunctional Landscapes Volume I: Theory, Values and History (Advances in Ecological Sciences)*. Southampton: WIT Press.

Bünemann, E.K., Bongiorno, G., Bai, Z., Creamer, R.E., De Deyn, G., De Goede, R., Fleskens, L., Geissen, V., Kuyper, T.W., Mäder, P., Pulleman, M., Sukkel, W., Van Groenigen, J., Brussaard, L. (2018) 'Soil quality – A critical review'. *Soil Biology and Biochemistry*, 120. pp.105-125.

Burns, R.G. (1978) 'Enzymes activity in soil: some theoretical and practical considerations', in *Soil Enzymes*. London: Academic Press, pp.295-340.

Burns, R.G. (1982) 'Enzyme activity in soil: Location and a possible role in microbial ecology'. *Soil Biology and Biochemistry,* 14(5). pp.423-427.

Burns, R.G., DaForest, J.L., Marxsen, J., Sinsabaugh, R.L., Stromberger, M.E., Wallenstein, M.D., Weintraub, M.N., Zoppini, A. (2013) 'Soil enzymes in a changing environment: Current knowledge and future directions'. *Soil Biology and Biochemistr*y, 58. pp.216-234.

Cardoso, E.J.B.N., Vasconcellos, R.L.F., Bini, D., Miyauchi, M.Y.H., de Santos, C.A., Alves, P.R.L., de Paula, A.M., Nakatani, A.S., Pereira, J.M., Nogueira, M.A. (2013) 'Soil health: looking for suitable indicators. What should be considered to assess the effects of use and management on soil health?'. *Scientia Agricola,* 70(4). pp.274-289.

Catt, J. A., Weir, A. H., Norrish, R. E., Rayner, J. H., King, D. W., Hall, D. G. M., Murphy, C. P. (1980) *The soils of Woburn Experimental Farm. III. Stackyard, Rothamsted Experimental Station Report for 1979, Part 2*. Harpenden: Rothamsted Research, pp.5-39.

Chen, J., Shen, C., Liu, C. (2010) 'N-Acetylglucosamine: Production and Applications'. *Marine Drugs,* 8(9). pp.2493-2516.

Chepil, W.S. (1962) 'A compact rotary sieve and the importance of dry sieving in physical soil analysis'. *Soil Science Society of America Proceedings*, 26(1). pp.4–6.

Cherubin, M.R., Karlen, D.L., Cerri, C.E., Franco, A.L., Tormena, C.A., Davies, C.A., Cerri, C.C. (2016) 'Soil quality indexing strategies for evaluating sugarcane expansion in Brazil'. *PLOS One*, 11(3). pp.1-26.

Clayden, B., Hollis, J. M. (1984). *Criteria for Differentiating Soil Series. Technical Monograph 17*. Harpenden: Soil Survey of England and Wales, pp.1-159.

Constanza, R., d'Arge, R., de Groots, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., Vandenbelt, M. (1997)

‘The value of the world’s ecosystem services and natural capital’. *Nature*, 387. pp.253-266.

Crossa, M.J., van Eeuwijk, F.A., Ramírez, M.E., Sayre, K. (1999) 'Using Partial Least Squares Regression, Factorial Regression, and AMMI Models for Interpreting Genotype × Environment Interaction'. *Crop Science*, 39(4). pp.955-967.

Crow, S.E., Hubanks, H., Deenik, J.L., Maaz, T.M., Glazer, C.T., Vizka, E., Riviera-Zayas, J. (2022) 'The legacy of intensive agricultural history on the soil health of (sub)tropical landscapes'. *Frontiers in Environmental Science*, 10. pp.1-13.

Curci, M., Pizzigallo, M., Crecchio, C., Mininni, R., Ruggiero, P. (1997) ‘Effects of conventional tillage on biochemical properties of soils’. *Biology and Fertility of Soils,* 25. pp.1-6.

DaForest, J.L. (2009) ‘The influence of time, storage temperature, and substrate age on potential soil enzyme activity in acidic forest soils using MUB-linked substrates and L-DOPA’ *Soil Biology and Biochemistry,* 41(6). pp.1180-1186.

DEFRA (2020) 25-Year Environment Plan. Available at: https://www.gov.uk/government/publications/25-year-environment-plan (Accessed September 21 2023)

Deng, S., Dick, R., Freeman, C., Kandeler, E., Weintraub, M.N. (2017) ‘Comparison and standardization of soil enzyme assay for meaningful data interpretation’ *Journal of Microbiological Methods,* 133. pp.32-34.

Devau, N., Le Cadre, E., Hinsinger, P., Jaillard, B., Gérard, F. (2009) 'Soil pH controls the environmental availability of phosphorus: Experimental and mechanistic modelling approaches'. *Applied Geochemistry,* 24(11). pp.2163-2174.

Dexter, A.R., Richard, G., Arrouays, D., Czyz, E.A., Jolivet, C., Duval, O. (2008) 'Complexed organic matter controls soil physical properties'. *Geoderma*, 144(3–4). pp.620–627.

Dick, R.P., Breakwell, D.P., Turco, R.F. (1997) 'Chapter 15: Soil Enzyme Activities and Biodiversity Measurements as Integrative Microbiological Indicators', in (eds) Doran, J.W., Jones, A.J. *Methods for Assessing Soil Quality,* 49. pp.247-271.

Dick, R.P. (1994). Soil enzyme activities as indicators of soil quality. In: Doran, J.W., Coleman, D.C., Bezdicek, D.F., Stewart, B.A. (Eds.), Defining Soil Quality for a Sustainable Environment. *Soil Science Society of America*, pp. 107-124.

Dodor, D. (2002) 'Enzyme activities in soils as affected by long-term cropping systems'. PhD Thesis, Iowa State University. Available at: https://dr.lib.iastate.edu/entities/publication/f1414562-1614-426a-811c-ab90d698e9e3. (Accessed 29 April 2022).

Doran, J.W., Parkin, T.B. (1994) 'Defining and Assessing Soil Quality', in Doran, J.W., Coleman, D.C., Bezdicek, D.F., Stewart, B.A. (eds.) *Defining Soil Quality for a Sustainable Environment*. Madison: Soil Science Society of America Journal, pp.3-21.

Doran, J.W., Zeiss, M.R. (2000) 'Soil health and sustainability: managing the biotic component of soil quality'. *Applied Soil Ecology*, 15(1). pp.3–11.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osbourne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S. (2013) 'Collinearity: a review of methods to deal with it and a simulation study evaluating their performance'. *Ecography*, 36. pp.27-46.

Dumas, J.B.A. (1831) 'Procedes de l'analyse Organic. Annales de Chimie et de Physique'. *Annals of Chemistry and of Physics,* 247. pp.198-213.

Dungait, J.A.J., Hopkins, D.W., Gregory, A.S., Whitmore, A.P. (2012) ‘Soil organic matter turnover is governed by accessibility not recalcitrance’. *Global Change Biology,* 18. pp.1781-1796.

Dunn, C., Jones, T.G., Girard, A., Freeman, C. (2014) ‘Methodologies for Extracellular Enzyme Assays from Wetland Soils’. *Wetlands,* 34. pp.9-17.

e-RA Rothamsted Archive (2023) Rothamsted Ley Arable: Highfield. Available at: https://www.era.rothamsted.ac.uk/experiment/rrn1 (Accessed December 09 2023).

E, Amézketa. (1999) ‘Soil Aggregate Stability: A Review’. *Journal of Sustainable Agriculture*, 14(2-3). pp.83-151.

Edmeades, D.C. (2003) 'The long-term effects of manures and fertilisers on soil productivity and quality: a review'. *Nutrient Cycling in Agroecosystems,* 66. pp.165-180.

Ekenler, M., Tabatabai, M.A. (2003) 'Responses of phosphatases and arylsulfatase in soils to liming and tillage systems'. *Journal of Plant Nutrition and Soil Science,* 166(3). pp.281-290.

Ekenler, M., Tabatabai, M. (2002) 'β-Glucosaminidase activity of soils: effect of cropping systems and its relationship to nitrogen mineralization'. *Biology and Fertility of Soils,* 36. pp.367-376.

Ellert, B.H., Clapperton, M.J., Anderson, D.W. (1997) ‘Chapter 5: An ecosystem perspective of soil quality’, in Gregorich, E.G., Carter, M.R. (ed.) *Developments in Soil Science*, 25. London: *Elsevier*, pp.115-141.

Environment Agency (2020) Environment Agency annual report and accounts 2019 to 2020. Available at: https://www.gov.uk/government/publications/environment-agency-annual-report-and-accounts-2019-to-2020. (Accessed June 12 2021)

Environment, Food and Rural Affairs Committee (2023) First Report of Session 2023–24. Available at: https://publications.parliament.uk/pa/cm5804/cmselect/cmenvfru/245/report.html. (Accessed November 14 2023)

European Commission (2023) Soil Health. Available at: https://environment.ec.europa.eu/topics/soil-and-land/soil-health_en#publications. (Accessed December 23 2023)

European Environment Agency (2024) CICES Version 5.1. Available at: https://cices.eu/ (Accessed January 23 2024)

Fajardo, M., McBratney, A. (2019) Slakes: A soil aggregate stability smart-phone app [Mobile application software]. Available at: https://play.google.com/store/apps/details?id=slaker.sydneyuni.au.com.slaker&hl=en. (Accessed 19 March 2023)

Feeney, J. (2019) 'Hunter-gatherer land management in the human break from ecological sustainability'. *The Anthropocene Review,* 6(3). pp.223-242

Fine, A.L., van Es, H.M., Schindelbeck, R.R. (2017) 'Statistics, scoring functions, and regional analysis of a comprehensive soil health database'. *Soil Science Society of America Journal*, 81(3). pp.589-601.

Fisher, B., Turner, R.K., Zylstra, M., Brouwer, R., de Groot, R., Farber, S., Ferraro, P., Green, R., Hadley, D., Harlow, J., Jefferiss, P., Kirkby, C., Morling, P., Mowatt, S., Naidoo, R., Paavola, J., Strassburg, B., Yu, D., Balmford, A. (2008) 'Ecosystem services

and economic theory: Integration for policy-relevant research'. *Ecological Applications,* 18. pp.2050-2067.

Flynn, K., Bagnall, D., Morgan, C. (2020) 'Evaluation of SLAKES, a smartphone application for quantifying aggregate stability, in high-clay soils'. *Soil Science Society of America*, 84(2). pp.345-353.

Food and Agriculture Organization of the United Nations (2003) *Chapter 4. Practices that influence the amount of organic matter*. Available at: https://www.fao.org/3/a0100e/a0100e07.htm#bm07l. (Accessed 11 November 2022)

Food and Agriculture Organization of the United Nations (2020) Land use in agriculture by the numbers. Available at: https://www.fao.org/sustainability/news/detail/en/c/1274219/ (Accessed December 14 2023)

Fowler, A., Basso, B., Maureira, F., Millar, N., Ulbrich. R., Brinton, W.F. (2024) 'Spatial patterns of historical crop yields reveal soil health attributes in US Midwest fields'. *Scientific Reports*, 14. pp.1-15.

Franson, M. (1980). *Standard Methods for the Examination of Water and Wastewater.* Washington D.C: American Public Health Association.

Gama-Rodrigues, A.C., Sales, M.V.S., Silva, P.S.D., Comerford, N.B., Cropper, W.P., Gama-Rodrigues, E.F. (2014) 'An exploratory analysis of phosphorus transformations in tropical soils using structural equation modeling'. *Biogeochemistry*, 118. pp.453-469.

German, D.P., Weintraub, M.N., Grandy, A.S., Lauber, C.L., Rinkes, Z.L., Allison, S.D. (2011) 'Optimization of hydrolytic and oxidative enzyme methods for ecosystem studies'. *Soil Biology and Biochemistry*, 43(7). pp.1387-1397.

Gomez, E.J., Delgado, J.A., Gonzalez, J.M. (2020) 'Environmental factors affect the response of microbial extracellular enzyme activity in soils when determined as a function of water availability and temperature'. *Ecology and Evolution,* 10(18). pp.10105-10115.

Government of India: Ministry of Agriculture and Farmers Welfare (2023) Soil Health Card. Available at: https://soilhealth.dac.gov.in/home (Accessed December 11 2023).

Gregorich, E.G., Janzen, H.H. (2000). Soil organic matter and agroecosystems. In: Sumner, M.E. (Ed.), *Handbook of Soil Science*. CRC Press.

Gugino, B.K., Idowu, O.J., Schindelbeck, R.R., van Es, H.M., Moebius-Clune, B.N., Wolfe, D.W., Thies, J.E., Abawi, G.S. (2009) '*Cornell Soil Health Assessment Training Manual, Edition 2*'. Geneva: Cornell University.

Guo, P., Wang, C., Feng, X., Su, M., Zhu, W., Tian, X. (2011) 'Mixed Inorganic and Organic Nitrogen Addition Enhanced Extracellular Enzymatic Activities in a Subtropical Forest Soil in East China'. *Water, Air, & Soil Pollution,* 216. pp.229-237.

Haghighi, F., Shorafa, M. (2010) 'A study of the effects of land use changes on soil physical properties and organic matter'. *Land Degradation & Development*, 21(5). pp.496-502.

Haney, L.R. (2012) '*Soil Health Test*'. Available at: https://www.midwestlabs.com/wp-content/uploads/ 2012/09/190Haney-Test-Explanation.pdf. (Accessed 12 June 2023)

Hengl, T., Leenaars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fergaus, E., Wheeler, I., Kwabena, N.A. (2017) 'Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning'. *Nutrient Cycling in Agroecosystems,* 109. pp.77-102.

Hernández, D.L., Hobbie, S.E. (2010) 'The effects of substrate composition, quantity, and diversity on microbial activity'. *Plant and Soil*, 335. pp.397–411.

Himmel, D. (1992) *Out of the Earth. Civilization and the Life of the Soil*. London: Aurum Press. pp.321.

Holmbery, J., Stephen, B., Lloyd, T. (1991). *Defending the future: a guide to sustainable development*. London: *Earthscan/IIED*, pp.13.

International Organization for Standardization (2012) Soil Quality – Measurement of the Stability of Soil Aggregates Subjected to the Action of Water. Available at: https://www.iso.org/obp/ui/#iso:std:iso:10930:ed-1:v1:en. (Accessed 06 October 2020)

IUSS Working Group WRB (2022) World Reference Base for Soil Resources. International soil classification system for naming soils and creating legends for soil maps. 4th edition. Available at: https://wrb.isric.org/files/WRB_fourth_edition_2022-12-18.pdf. (Accessed 14 October 2023)

Jatish, C., Naher, U.A. (2019) 'Chapter 21 - Soil Nutrient Stress and Rice Production in Bangladesh', in Hasanuzzaman, M., Fujita, M., Nahar, K., Biswas, J.B. (eds.) *Advances in Rice Research for Abiotic Stress Tolerance*. Duxford: Woodhead Publishing, pp.431-445.

Jenkinson, D.S. (1991) 'The Rothamsted Long-term Experiments: are they still of use?'. *Agronomy Journal*, 83. pp.2-10.

Johnston, A. E. (1972) 'The effects of ley and arable cropping systems on the amounts of soil organic matter in the Rothamsted and Woburn ley-arable experiment'. *Rothamsted Experimental Station Report*, 2. pp.131-159.

Johnston, A.E., Poulton, P.R., Macdonald, A.J., Glendining, M. J. (2022) 'The Woburn Ley-arable experiment, 1938-2020: its management, the issues and problems that arose, and their resolution to maintain the continuity and relevance of the experiment'. Harpenden: Rothamsted Research, pp.1-28.

Jónsson, J.G., Davíðsdóttir, B. (2016) 'Classification and valuation of soil ecosystem services'. *Agricultural Systems,*145. pp.24-38.

Kalembasa, S., Symanowicz, B. (2012) 'Enzymatic Activity of Soil after Applying Various Waste Organic Materials, Ash, and Mineral Fertilizers'. *Polish Journal of Environmental Studies,* 21. pp.1635-1641.

Kallenbach, C.M., Frey, S.D., Grandy, A.S. (2016) 'Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls'. *Nature Communications*, 7. pp.1-10.

Karlen, D.L., Andrews, S.S., Doran J.W. (2001) 'Soil quality: Current concepts and applications'. *Advances in Agronomy,* 74. pp.1-40.

Karlen, D.L., Ditzler, C.A., Andrews, S.S.(2003) 'Soil quality: why and how?'. Geoderma, 114. pp.145–156.

Karlen, D.L., Mausbach, M.J., Doran, J.W., Cline, R.G., Harris, R.F., Schuman, G.E. (1997) 'Soil quality: a concept, definition and framework for evaluation'. *Soil Science Society of America Journal,* 61. pp.4-10.

Karlen, L.D., Stott, D.E. (1994) 'A Framework for Evaluating Physical and Chemical Indicators of Soil Quality', in Doran, J.W., Coleman, D.C., Bezdicek, D.F., Stewart, B.A. (eds.) *Defining Soil Quality for a Sustainable Environment*. Madison: *Soil Science Society of America Journal*, pp.53-72.

Keller, N., Bol, R., Herre, M., Marschner, B., Heinze, S. (2023) ‘Catchment scale spatial distribution of soil enzyme activities in a mountainous German coniferous forest’. *Soil Biology and Biochemisty*, 177(1). pp.108885.

Kemper, W.D., Rosenau, R.C. (1986) ‘Aggregate Stability and Size Distribution’, in Klute, A. (ed.) *Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods, 2nd Edition*. American Society of Agronomy, Soil Science Society of America. Madison, WI: USA, pp.425 - 442.

Khokhar, A., Kusuma Grace, J., Srinivasa-Rao, C., Maruthi-Sankar, G.R., Ravindrachary, G., Sammi, R.K., Srinivas, K., Lal, M., Satish, K.T., Usha Rani, K. (2014) ‘Combined effect of tillage and organic fertilization on soil quality key indicators and indices in alluvial soils of Indo-Gangetic Plains under rainfed maize–wheat system’. *Archives of Agronomy and Soil Science*, 61(3). pp.313-327.

Kibblewhite, M.G., Ritz, K., Swift, M.J. (2007) ‘Soil health in agricultural systems’. *Philosophical Transactions of the Royal Society B,* 363(1492). pp.685-701.

Kim, M., Lee, S., Park, H., Kim, J. (2021) 'Evaluation of Two Amendments (Biochar and Acid Mine Drainage Sludge) on Arsenic Contaminated Soil Using Chemical, Biological, and Ecological Assessments'. *Materials,* 14(15). pp.4111.

Klingebiel, A. and Montgomery, P.H. (1961) *Land-capability classification*. Agricultural Handbook No. 210. Washington, DC: US Department of Agriculture Soil Conservation Service.

Knight, T.R., Dick, R.P. (2004) ‘Differentiating microbial and stabilized β-glucosidase activity relative to soil quality’. *Soil Biology and Biochemistry,* 36(12). pp.2089-2096.

Koch, A., McBratney, A., Adams, M., Field, D., Hill, R., Crawford, J., Minasny, B., Lal, R., Abbott, L., O’Donnell, A., Angers, D., Baldock, J., Barbier, E., Binkley, D., Parton, W., Wall, D.H., Bird, M., Bouma, J., Chenu, C., Flora, C.B., Goulding, K., Grunwald, S.,

Hempel, J., Jastrow, J., Lehmann, J., Lorenz, K., Morgan, C.L., Rice, C.W., Whitehead, D., Young, I., Zimmermann, M. (2013) 'Soil Security: Solving the Global Soil Crisis'. *Global Policy,* 4(4). pp.434-441.

Kopittke, P.M., Menzies, N.W., Wang, P., McKenna, B.A., Lombi, E. (2019) 'Soil and the intensification of agriculture for global food security'. Environment International, 132. pp.105078.

Kuipers, H. (1991) 'Agronomic aspects of ploughing and non-ploughing links open overlay panel'. *Soil and Tillage Research*, 21(1-2). pp.167-176.

Kunlanit, B., Khwanchum, L., Vityakon, P. (2020) 'Land Use Changes Affecting Soil Organic Matter Accumulation in Topsoil and Subsoil in Northeast Thailand'. *Applied and Environmental Soil Science,* 2020(1). pp.1-15.

Lal, R. (2001) 'Managing world soils for food security and environmental quality' *Advances in Agronomy,* 74. pp.155-192.

Lal, R. (2016) 'Soil health and carbon management'. *Food and Energy Security,* 5(3). pp.212–222.

Lal. R. (2004) 'Soil Carbon Sequestration Impacts on Global Climate Change and Food Security'. *Science,* 304. pp.1623-1627.

Le Bissonnais, Y. (2016) 'Aggregate stability and assessment of soil crustability and erodibility: I. Theory and methodology'. *European Journal of Soil Science*, 67. pp.1-21.

Lee, M., Shin, A., Oh, J., Hwang, D., Ko, Y., Hyeon, J., Han, S. (2021) 'Enzymatic production of sugar from fungi and fungi-infected lignocellulosic biomass by a new cellulosomal enzyme harboring N-acetyl-β-d-glucosaminidase activity'. *Bioresource Technology,* 319. pp.124242.

Lee, Y., Lorenz, N., Dick, L.K., Dick, R.P. (2007) ‘Cold Storage and Pretreatment Incubation Effects on Soil Microbial Properties’. *Soil Science Society of America Journal,* 71(4). pp.1299-1305.

Lehmann, J., Bossio, D.A., Kögel-Knabner, I., Rillig, M.C. (2020) 'The concept and future prospects of soil health'. *Nature Reviews Earth & Environment*, 1. pp.544-553.

Lellei-Kovács, E., Kovács-Láng, E., Botta-Dukat, Z., Kalapos, T., Emmett, B., Beier, C. (2011) ‘Thresholds and interactive effects of soil moisture on the temperature response of soil respiration’. *European Journal of Soil Biology,* 47. pp.247-255.

Li, X., Jia, B., Lv, J., Ma, Q., Kuzyakov, Y., Li, F. (2017) 'Nitrogen fertilization decreases the decomposition of soil organic matter and plant residues in planted soils'. *Soil Biology and Biochemistry,* 112. pp.47-55.

Lilburne, L., Sparling, G., Schipper, L. (2004) ‘Soil quality monitoring in New Zealand: development of an interpretative framework’. *Agriculture, Ecosystems & Environment*, 104(3). pp.535-544.

Lv, L., Gao, Z., Liao, K., Zhu, Q., Zhu, J. (2023) 'Impact of conservation tillage on the distribution of soil nutrients with depth'. *Soil and Tillage Research,* 225. pp.105527.

Maaz, T., Brown, K, Smith, P, and Johnson, J. (2023). Applying structural equation modeling to disentangle the effects of soil properties, management practices, and ecosystem functions. *Journal of Soil Science and Ecology*, 12(4), pp. 456-471.

Maaz, T.M., Heck, R.H., Glazer, C.T., Loo, M.K., Riviera-Zayas, J., Krenz, A., Beckstrom, T., Crow, S.E., Deenik, J.L. (2023) 'Measuring the immeasurable: A structural equation modeling approach to assessing soil health'. *Science of the Total Environment*, 870. pp.161900.

Madejón, E., Moreno, F., Murillo, J.M., Pelegrín, F. (2007) 'Soil biochemical response to long-term conservation tillage under semi-arid Mediterranean conditions'. *Soil and Tillage Research,* 94(2). pp.346-352.

Maes, M.J.A., Jones, K.E., Toledano, M.B., Milligan, B. (2020) 'Accounting for natural capital has cross-cutting relevance for UK public sector decision-making'. *Ecosystem Services,* 44. pp.101127.

Marchant, B.P. (2021) 'Using remote sensors to predict soil properties: Radiometry and peat depth in Dartmoor, UK'. *Geoderma,* 403. pp.115232.

Marx, M.C., Wood, M., Jarvis, S.C. (2001) 'A microplate fluorimetric assay for the study of enzyme diversity in soils'. *Soil Biology and Biochemistry,* 33(12-13). pp.1633-1640.

Mausel, P.W. (1971) 'Soil quality in Illinois—an example of a soils geography resource analysis'. *The Professional Geographer,* 23. pp.127-136.

Mclean, E.O. (1983) 'Chapter 12: Soil pH and lime requirement', in Page, A.L. (ed.) *Methods of Soil Analysis: Part 2 Chemical and Microbiological Properties (2nd edition)*. Wisconsin: *American Society of Agronomy and Soil Sciences Society of America*, pp.199-225.

Millennium Ecosystem Assessment (2005) *Millennium Ecosystem Assessment: Current State and Trends Assessment*. Washington D.C: Island Press.

Moncada, M.P., Gabriels, D., Cornelis, W.M. (2014) 'Data-driven analysis of soil quality indicators using limited data'. *Geoderma*, 235. pp.271–278.

Monokrousos, N., Papatheodorou, E.M., Stamou, G.P. (2008) 'The response of soil biochemical variables to organic and conventional cultivation of Asparagus sp.'. *Soil Biology and Biochemistry,* 40(1). pp.198-206.

Monreal, C.M., Bergstrom, D.W. (2000) 'Soil enzymatic factors expressing the influence of land use, tillage system and texture on soil biochemical quality'. *Canadian Journal of Soil Science,* 80(3). pp.419-428.

Morrow, J.G., Huggins, D.R., Carpenter-Boggs, L.A., Reganold, J.P. (2016) 'Evaluating measures to assess soil health in long-term agroecosystem trials'. *Soil Science Society of America Journal*, 80. pp.450-462.

Muthukrishnan, S., Merzendorfer, S., Arakane, Y., Kramer, K.J. (2012) '7 - Chitin Metabolism in Insects', in Lawrence, G.I. (ed.) *Insect Molecular Biology and Biochemistry*. London: Academic Press, pp.193-235.

Nakayama, Y., Wade, J., Li, C., Daughtridge, R.C., Margenot, A.J. (2023) 'Quantifying the relative importance of controls and assay conditions for reliable measurement of soil enzyme activities with para-nitrophenol substrates'. *Geoderma,* 429. pp.116234.

National Research Council (1993) *Soil and Water Quality: An Agenda for Agriculture*. Washington, DC: The National Academies Press.

Neal, A. L., Bacq-Labreuil, A., Zhang, X., Clark, I.M., Coleman, K., Mooney, S.J., Ritz, K., Crawford, J. W. (2020) 'Soil as an Extended Composite Phenotype of the Microbial Metagenome'. *Scientific Reports*, 10. pp.10649.

Nielsen, M.N., Winding, A. (2002) 'Microorganisms as indicators of soil health'. *National Environmental Research Institute, Denmark Technical Report,* 388. pp.7-64.

Noponen, M.R.A., Healey, J.R., Soto, G., Haggar, J.P. (2013) 'Sink or source—The potential of coffee agroforestry systems to sequester atmospheric CO2 into soil organic carbon'. *Agriculture, Ecosystems & Environment*, 175. pp.60-68.

North Carolina State University (2023) Soil Acidity and Liming: Basic Information for Farmers and Gardeners.Available at: https://content.ces.ncsu.edu/soil-acidity-and-liming-basic-information-for-farmers-and-gardeners. (Accessed 19 October 2023)

O'Farrell, P.J., Anderson, P.M.L. (2010) 'Sustainable multifunctional landscapes: a review to implementation'. *Current Opinion in Environmental Sustainability,* 2. pp.59-65.

Obade, V.P., Lal, R. (2016) 'Towards a standard technique for soil quality assessment'. *Geoderma*, 265. pp.96-102.

Omonode, R.A., Vyn, T.J. (2006) 'Vertical distribution of soil organic carbon and nitrogen under warm-season native grasses relative to croplands in west-central Indiana, USA'. *Agriculture, Ecosystems & Environment*, 117(2-3). pp.159-170.

Pahalvi, H.N., Rafiya, L., Rashid, S., Nisar, B., Kamili, A.N. (2021) 'Chemical Fertilizers and Their Impact on Soil Health', in Dar, G.H., Bhat, R.A., Mehmood, M.A., Hakeem, K.R. (eds) *Microbiota and Biofertilizers, 2.* Cham: Springer, pp.1-20.

Pandey, D., Agrawal, M., Bohra, J.S. (2014) 'Effects of conventional tillage and no tillage permutations on extracellular soil enzyme activities and microbial biomass under rice cultivation'. *Soil and Tillage Research,* 136. pp.51-60.

Parham, J.A., Deng, S.P. (2000) 'Detection, quantification and characterization of β-glucosaminidase activity in soil'. *Soil and Biochemistry,* 32(8-9). pp.1183-1190.

Patzel, N., Sticher, H., Karlen, D.L. (2000) 'Soil Fertility — Phenomenon and Concept'. *Journal of Plant Nutrition and Soil Science,* 163(2). pp.129-142.

Pereira, P., Bogunovic, I., Muñoz-Rojas, M., Brevik, E.C. (2018) 'Soil ecosystem services, sustainability, valuation and management'. *Current Opinion in Environmental Science & Health,* 5. pp.7-13.

Perring, M.P., De Frenne, P., Baeten, L., Maes, S.L., Depauw, L., Blondeel, H., Carón, M.M. and Verheyen, K. (2016) 'Global environmental change effects on ecosystems: the importance of land-use legacies'. *Global Change Biolology,* 22. pp.1361-1371.

Piotrowska-Długosz, A., Kobierski, M., Długosz, J. (2021) 'Enzymatic Activity and Physicochemical Properties of Soil Profiles of Luvisols'. *Materials,* 14(21). pp.6364.

Piotrowska-Długosz, A., Wilczewski, E. (2014) 'Changes in enzyme activities as affected by green-manure catch crops and mineral nitrogen fertilization'. *Zemdirbyste,* 101. pp.139-146.

Portnoy, S. (1988) 'Asymptotic Behavior of Likelihood Methods for Exponential Families when the Number of Parameters Tends to Infinity'. *The Annals of Statistics*, 16(1). pp.356-366.

Powlson, D. (2020) Soil health—useful terminology for communication or meaningless concept? Or both? Frontier Agriculture Science Engineering 7(3): pp.246–250 https://doi.org/10.15302/J-FASE-2020326

Redmile-Gordon, M., Gregory, A.S., White, R.P., Watts, C.W. (2020) 'Soil organic carbon, extracellular polymeric substances (EPS), and soil structural stability as affected by previous and current land-use'. *Geoderma*, 363. pp.114143.

Ribeiro-Kumara, C., Pumpanen, J., Heinonsalo, J., Metslaid, M., Orumaa, A., Jõgiste, K., Berninger, F., Köster, K. (2020) 'Long-term effects of forest fires on soil greenhouse gas emissions and extracellular enzyme activities in a hemiboreal forest'. *Science of The Total Environment*, 718. pp.135291.

Rieke, E.L., Bagnall, D.K., Morgan, C.L.S., Flynn, K.D., Howe, J.A., Greub, K.L.H., Mac Bean, G., Cappellazzi, S.B., Cope, M., Liptzin, D., Norris, C.E., Tracy, P.W., Aberle, E., Ashworth, A., Tavarez, O.B., Bary, A.I., Baumhardt, R.L., Gracia, A.B., Brainard, D.C.,

Brennan, J.R., Honeycutt, C.W. (2022) ‘Evaluation of aggregate stability methods for soil health’. Geoderma, 428. pp.116-156.

Riggs, C.E., Hobbie, S.E., Bach, E.M., Hofmockel, K.S., Kazanski, C.E. (2015) ‘Nitrogen addition changes grassland soil organic matter decomposition’. *Biogeochemistry,* 125. pp.203-219.

Ritz, K., Black, H.I., Campbell, C.D., Harris, J.A., Wood, C. (2009) ‘Selecting biological indicators for monitoring soils: a framework for balancing scientific and technical opinion to assist policy development’. *Ecological Indicators*, 9(6). pp.1212-1221.

Romig, D.E., Garylynd, M.J., Harris, R.F., McSweeney, K. (1995) ‘How farmers assess soil health and quality’. *Journal of Soil and Water Conservation,* 50(3). pp.1-229.

Roper, W.R., Osmond, D.L., Heitman, J.L. (2019) 'A Response to “Reanalysis Validates Soil Health Indicator Sensitivity and Correlation with Long-term Crop Yields”'. *Soil Science Society of America Journal*, 83(6). pp.1842-1845.

Rosseel, Y. (2011) 'lavaan: An R Package for Structural Equation Modeling'. *Journal of Statistical Software*, 48(2). pp.1-36.

Saha, S., Mina, B.L., Gopinath, K.A., Kundu, S., Gupta, H.S. (2007) 'Organic amendments affect biochemical properties of a subtemperate soil of the Indian Himalayas'. *Nutrient Cycling in Agroecosystems,* 80. pp.233-242.

Saiya-Cork, K.R., Sinsabaugh, R.L., Zak, D.R. (2002) 'The effects of long term nitrogen deposition on extracellular enzyme activity in an Acer saccharum forest soil'. *Soil Biology and Biochemistry*, 34(9). pp.1309-1315.

Sardans, J., Peñuelas, J., Ogaya, R. (2008) 'Experimental drought reduced acid and alkaline phosphatase activity and increased organic extractable P in soil in a Quercus ilex Mediterranean forest'. *European Journal of Soil Biology,*44(5-6). pp.509-520.

Saviozzi, A., Bufalino, P., Levi-Minzi, R., Riffaldi, R. (2002) 'Biochemical activities in a degraded soil restored by two amendments: a laboratory study'. *Biology and Feritlity of Soils,* 35. pp.96-101.

Saviozzi, A., Levi-Minzi, R., Cardelli, R., Riffaldi, R. (2001) 'A comparison of soil quality in adjacent cultivated, forest and native grassland soils'. *Plant and Soil*, 233. pp.251-259.

Schaeffer, A. (2021) 'Pesticide Effects on Enzyme Activities in the Soil Ecosystem'. *Soil Biochemistry,* 8. pp.273-340.

Schimel, J., Bennett, J., Fierer, N. (2005) '10 -Microbial community composition and soil nitrogen cycling: is there really a connection?', in Bardgett, R., Usher, M., Hopkins, D. (eds.) *Biological Diversity and Function in Soils*. Cambridge: Cambridge University Press, pp.171-188.

Schimel, J.P., and Bennett, J. (2004). Nitrogen mineralization: challenges of a changing paradigm. *Ecology*, 85(3), pp.591-602.

Schindelbeck, B., Kurtz, K. (2017) 'Soil Health Manual Series - Fact Sheet Number 16-07: Wet Aggregate Stability'. Cornell University School of Integrative Plant Sciences, Cornell University Soil Health Laboratory. pp.1-2. Available at: https://bpb-us-e1.wpmucdn.com/blogs.cornell.edu/dist/7/9922/files/2021/11/07_AgStab.pdf. (Accessed 12 June 2020)

Schulte, R.P.O., Bampa, F., Bardy, M. (2006) 'A critical review of the current soil quality assessment models and indicators'

Shi, W. (2010) 'Agricultural and Ecological Significance of Soil Enzymes: Soil Carbon Sequestration and Nutrient Cycling'. *Soil Enzymology,* 22. pp.43-60.

Shi, W., Yao, J., Yan, F. (2009) ‘Vegetable cultivation under greenhouse conditions leads to rapid accumulation of nutrients, acidification and salinity of soils and groundwater contamination in South-Eastern China’. *Nutrient Cycling in Agroecosystems,* 83. pp.73-84.

Shukla, G., Varma, A. (2010) ‘Chapter 2: Role of enzymes in maintaining soil health’, in *Soil Enzymology*. Berlin:Springer. pp.25-43.

Singh, P.K., Singh, M., Tripathi, B.N. (2012) ‘Glomalin: an arbuscular mycorrhizal fungal soil protein’. *Protoplasma,*250. pp.663-669.

Six, J., Elliott, E.T., Paustian, K. (2000) 'Soil macroaggregate turnover and microaggregate formation: a mechanism for C sequestration under no-tillage agriculture'. *Soil Biology and Biochemistry,* 32(14). pp.2099-2103.

Sojka, R.E., Upchurch, D.R. (1999) ‘Reservations regarding the soil quality concept’. *Soil Science Society of America Journal,* 63. pp.1039-1054.

Sojka, R.E., Upchurch, D.R., Borlaug, N.E. (2003) ‘Quality soil management or soil quality management: performance versus semantics’. *Advances in Agronomy,* 79. pp.1-68.

Sotomayor-Ramírez, D., Espinoza, Y., Acosta-Martínez, V. (2009) 'Land use effects on microbial biomass C, β-glucosidase and β-glucosaminidase activities, and availability, storage, and age of organic C in soil'. *Biology and Fertility of Soils,* 45. pp.487-497.

Spark, K.M., Swift, R.S. (2002) ‘Effect of soil composition and dissolved organic matter on pesticide sorption’. *Science of The Total Environment,* 298(1-3). pp.147-161.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J. (2010) ‘Visible and Near Infrared Spectroscopy in Soil Science’, in Sparks, D.L. (ed.) *Advances in Agronomy, Vol. 107*. Burlington: Academic Press, pp.163-215.

Stockdale, E.A., Powlson, D.S. (2003) ‘Can we measure soil quality - and who cares?’ Unpublished discussion paper for Southeast Soils Discussion Group of British Society of Soil Science. Available at: https://repository.rothamsted.ac.uk/download/7673749a5f8edf8f55fca717aaad1091e134fbcb49214e2fa5db230d57690295/3557371/Soil%20health%20FASE%202020.pdf

Sun, R., Niu, J., Luo, B., Wang, X., Li, W., Zhang, W., Wang, F., Zhang, C., Ye, X. (2022) 'Substitution of manure for mineral P fertilizers increases P availability by enhancing microbial potential for organic P mineralization in greenhouse soil'. *Frontiers in Bioengineering and Biotechnology,* 10. pp.1078626

Svoray, T., Hassid, I., Atkinson, P.M., Moebius-Clune, B.N., van Es, H.M. (2015) ‘Mapping soil health over large agriculturally important areas’. *Soil Science Society of America Journal*, 79(5). pp.1420-1434.

Tahir, S., Marschner, P. (2016) 'Clay amendment to sandy soil—effect of clay concentration and ped size on nutrient dynamics after residue addition’. *Journal of Soils and Sediments,* 16. pp.2072-2080.

Teasdale, J.R., Coffman, C.B., Mangum, R.W. (2007) ‘Potential Long-Term Benefits of No-Tillage and Organic Cropping Systems for Grain Production and Soil Improvement’. *Agronomy Journal,* 99(5). pp.1297-1305.

Tisdall, J.M., and Oades, J.M. (1982). Organic matter and water-stable aggregates in soils. *Journal of Soil Science*, 33(2), pp.141-163.

Truman, C.C., Bradford, J.M., Ferris, J.E. (1990) 'Antecedent Water Content and Rainfall Energy Influence on Soil Aggregate Breakdown'. *Soil Science Society of America Journal*, 54(5). pp.1385-1392.

Turner, B.L., Hopkins, D.W., Haygarth, P.M., Ostle, N. (2002) ‘β-Glucosidase activity in pasture soils’. *Applied Soil Ecology,* 20(2). pp.157-162.

UN (2023) Land - the planet’s carbon sink. Available at: https://www.un.org/en/climatechange/science/climate-issues/land (Accessed December 12 2023).

Uzuner, S., Cekmecelioglu, D. (2019) ‘Chapter 3 - Enzymes in the Beverage Industry’, in Kuddus, M. (ed.) *Enzymes in Food Biotechnology*. London: Academic Press, pp.29-43.

van Es, H.M., Karlen, D.L. (2019) 'Reanalysis Validates Soil Health Indicator Sensitivity and Correlation with Long-term Crop Yields'. *Soil Science Society of America Journal*, 83(3). pp.721-732.

Veum, K.S., Sudduth, K.A., Kremer, R.J., Kitchen, N.R. (2015) ‘Estimating a Soil Quality Index with VNIR Reflectance Spectroscopy’. *Soil Science Society of America Journal*, 79. pp.637-649

Walinga, I., Van Vark, W., Houba, V.J.G., Van Der Lee, J.J. (1989) 'Plant Analysis Procedures'. *Department of Soil Science and Plant Nutrition, Wageningen Agricultural University,* 7. pp.197-200.

Wallenius, K., Rita, H., Simpanen, S., Mikkonen, A., Niemi, R.M. (2010) ‘Sample storage for soil enzyme activity and bacterial community profiles’. *Journal of Microbiological Methods,* 81(1). pp.48-55.

Wallenstein, M.D., Weintraub, M.N. (2008) 'Emerging tools for measuring and modeling the in situ activity of soil extracellular enzymes'. *Soil Biology and Biochemistry*, 40(9). pp.2098-2106.

Wang, X., Lu, Q. (2006) 'Beta-Glucosidase Activity in Paddy Soils of the Taihu Lake Region, China'. *Pedosphere,*16(1). pp.118-124.

Warkentin, B.P., Fletcher, H.F. (1977) *Proceedings of the International Seminar on Soil Environment and Fertility Management in Intensive Agriculture*. Tokyo: National Institute of Agricultural Science. pp.594-598.

Wu, Q., Dou. X., Wang, Q., Guan, Z., Cai, Y., Liao, X. (2018) 'Isolation of β-1,3-Glucanase-Producing Microorganisms from Poria cocos Cultivation Soil via Molecular Biology'. *Molecules,* 23(7). pp.1555.

Yan, F., Schubert, S., Mengel, K. (1996) 'Soil pH increase due to biological decarboxylation of organic anions'. *Soil Biology and Biochemistry*, 28(4-5). pp.617-624.

Zhang, Z., Tang, G., Chai, X., Liu, B., Gao, X., Zeng, F., Wang, Y., Zhang, B. (2023) 'Different Responses of Soil Bacterial and Fungal Communities in Three Typical Vegetations following Nitrogen Deposition in an Arid Desert'. *Microorganisms,* 11(10). pp.2471.

Zheng, Y., Han, X., Li, Y., Yang, J., Li, N., An, N. (2019) 'Effects of Biochar and Straw Application on the Physicochemical and Biological Properties of Paddy Soils in Northeast China'. *Scientific Reports,* 9(1). pp.16531.