Biochemical Soil Health Indicator Scores based on a Multivariate Soybean Yield Prediction Model

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Peter Renz

Graduate Program in Environment & Natural Resources

The Ohio State University

2024

Dissertation Committee:

Dr. Richard P. Dick, Advisor

Dr. Rattan Lal

Dr. Brian K. Slater

Dr. Laura E. Lindsey

Abstract

Soils are a non-renewable resource, which is the foundation of all ecosystems. Mismanagement of soil particularly in agro-ecosystems has degraded soil. To guide management of soils, to remediate soils, and enable optimal agricultural production, soil health indicators are needed.

The objective of this dissertation was to determine the potential of biological and other soil properties to predict soybean yields. The central approach was based on soil samples from farmers' fields instead of long-term experimental sites (LTES). Farmers' fields in this study represented diverse management practices that exist in the agricultural sector. Soil Health (SH) measurements that are calibrated and that can consistently detect land management are lacking which was shown in Roper's et al. (2017) 2017 publication that found existing SH tests (CASH, Haney) had limited ability to identify agronomic land management practices at a North Caroline LTES. And that they were poorly correlated with crop yields. This means that the quote by the Soil Health Institute "There is no standardized measurement for Soil Health in the United States" is still true.

Unlike most previous research on SH, which was based on data from long-term experimental sites (LTES), this investigation utilized analyses of soil samples from farmers' fields. Farmers were surveyed to collect historical management information on each field after which LIDAR data, and soil type information from the Soil Survey website was obtained. For the 2019, 2020, 2021 growing seasons Ohio sampling sites were visited during the spring season and a composite soil sample at a

depth of 0-15 cm was collected. In 2021 soil samples were collected at three LTES in Ohio and one in Michigan. Furthermore, soils were sampled at two virgin and two restored prairie sites in Ohio.

Soil samples were analyzed for microbial community composition, enzyme activities, total carbon (TC), soil organic carbon (SOC), total nitrogen (TN), pH and texture. Microbial communities were profiled using the Ester-Linked Fatty Acid Methyl Ester (EL-FAME) analysis. The enzyme activity of β-glucosidase (NAG), N-acetyl glutamate synthase (NAG), and arylsulfatase (AS) were chosen because previous research has shown these measurements have been shown to be sensitive in detecting soil/crop management effects.

The results of this dissertation showed that no individual soil property variable or other variable had a strong relationship with soybean yield. The multi-variate regression analysis on the other hand resulted in a correlation of determination value (R^2) of 0.84. In this analysis a statistical machine learning algorithm (Elastic Net) was used with the help of the *glmnet* R package. The optimized model was then used to develop the biochemical Soil Heath Index (SHI) by extracting the individual regression coefficients of all biological variables (e.g. enzyme and EL-FAME variables) and the usage of a mathematical algorithem. The most common SH indicators in this study and the computed SH scores were analyzed for their ability to detect soil management at four LTES by running the Tukey's Honest Significant Difference test in combination with a sensitivity scoring algorithm. The sensitivity scores were used to identify the most sensitive SH indicators.

In the final analysis 512 soil variables were scored for their ability to detect

agricultural land management practices (e.g. crop rotation, cover cropping, soil amendments, tillage practices), restored prairies, and virgin soil in Ohio. Additionally at the agricultural scale each variable was tested for its temporal sensitivity. The most sensitive SH indicators were identified with the help of a sensitivity scoring algorithm and their relationship to soil organic carbon was determined. The remaining SH indicators were used to determine beneficial and detrimental agricultural land management practices.

Dedication

This work is dedicated to Aaron Maurer, a family member that is dearly missed by everyone, my family in Germany and the USA, my lovely wife Laura, and our wonderful two children. Without their support and love this work would have been much more difficult to accomplish.

Acknowledgments

I am eternally thankful to the Ohio Soybean Council for funding my research and giving me the opportunity to work with Ohio farmers to explore a new approach to define the health of a soil in an agricultural framework.

I am also grateful for the following people:

Dr. Richard P. Dick, for trusting me in my abilities to conduct a research project of such a scale and provide me with advice and mentoring me. To Dr. Laura E. Lindsey for listening to my questions related to soybean land management practices and proving me with a survey foundation that was then adapted for the study. To my remaining committee members who provided their time and suggestions to me: Dr. Rattan Lal, and Dr. Brian K. Slater.

A large thank you goes to Dr. Nicola Lorenz and everyone that helped me in the lab. Nicola was an essential connecting point for me. She managed all the necessary equipment orders and provided me with information and suggestions that were very helpful and appreciated.

I also want to thank Nathan Lee, who trained me on the elemental analyzer, the gas chromatograph, and the enzyme activity assays. Without him the start into my research project would have taken some more time.

I want to thank all those friends who helped me sieve my wet soil samples through the 2 mm sieve throughout these three growing seasons. My fingers, hand, and wrist really appreciate the help.

Finally, I want to thank once again my parents and all the professors and teachers that educated me during my lifetime. Without their mentorship I would not be where I am today.

Curriculum Vitae

September 1984	Born - Dresden, Saxony, Germany
2005-2007	Material Science, Technical University Dresden, Germany
2007-2012	M. Eng. (Diplomingenieur), Waste Management and Contaminated Site Treatment, Technical University Dresden, Germany
2010-2011	Visiting Scholar, Dr. Makoto Nishigaki, Okayama University, Japan
2011-2012	Visiting Scholar, Dr. Richard P. Dick, The Ohio State University, USA
2019 to present	Graduate PhD Student and Teaching Assistant, The Ohio State University, USA

Fields of Study

Major Field: Environment and Natural Resources

Specialization: Soil Science

Table of Contents

List of Tables

List of Figures

Dissertation Introduction

Soils are a non-renewable resource, which is the foundation of all ecosystems. Mismanagement of soil particularly in agro-ecosystems has degraded soil. To guide management of soils, to remediate soils, and enable optimal agricultural production, soil health indicators are needed.

The objective of this dissertation was to determine the potential of biological and other soil properties to predict soybean yields. The central approach was based on soil samples from farmers' fields instead of long-term experimental sites (LTES). Farmers' fields in this study represented diverse management practices that exist in the agricultural sector. Soil Health (SH) measurements that are calibrated and that can consistently detect land management are lacking which was shown in Roper's et al. (2017) 2017 publication that found existing SH tests (CASH, Haney) had limited ability to identify agronomic land management practices at a North Caroline LTES. And that they were poorly correlated with crop yields. This means that the quote by the Soil Health Institute "There is no standardized measurement for Soil Health in the United States" is still true.

Extensive research has found certain soil enzyme assays to be quite sensitive for detecting land management effects and exhibit seasonal stability. The currently promoted SH indicator scores have limited or inappropriate biological indicators (e.g. microbial biomass and respiration). The latter measurements vary too much on a seasonal basis due to weather variation or a recent short term soil management event (e.g. high organic inputs, disturbance). Thus, the global objective of this dissertation

was to determine the potential of biological soil properties, specifically enzyme activities and microbial community biomarkers to quantify SH. Enzyme activity has the added advantage over most other soil biological measurements, that it can be run on air-dried soil and furthermore is relatively simple. This is attractive to commercial labs who want to minimize analytical costs and prefer to use air-died soils. A secondary objective was to determine the potential of soil properties to predict soybean yield with an individual variable or in a multivariate model.

Unlike most previous research on SH, which was based on data from long-term experimental sites (LTES), this investigation utilized analyses of soil samples from farmers' fields. Farmers were surveyed to collect historical management information on each field after which LIDAR data, and soil type information from the Soil Survey website was obtained. For the 2019, 2020, 2021 growing seasons Ohio sampling sites were visited during the spring season and a composite soil sample at a depth of 0-15 cm was collected. In 2021 soil samples were collected at three LTES in Ohio and one in Michigan. Furthermore, soils were sampled at two virgin and two restored prairie sites in Ohio.

In fall of each year, on farm fields that grew soybeans, soybean yields were determined at each soil sampling site. Soil samples were analyzed for microbial community composition, enzyme activities, total carbon (TC), soil organic carbon (SOC), total nitrogen (TN), pH and texture. Microbial communities were profiled using the Ester-Linked Fatty Acid Methyl Ester (EL-FAME) analysis. The enzyme activity of β-glucosidase (NAG), N-acetyl glutamate synthase (NAG), and arylsulfatase (AS) were chosen because previous research has shown these measurements have been shown to

be sensitive in detecting soil/crop management effects.

Chapter 1 of this dissertation reviewed the literature on SH.

The objective of Chapter 2 was to determine if enzyme activities, individual soil properties, or other variables could predict soybean yield. The study used soils from farm and long-term experimental sites (LTES) in Ohio and one LTES in Michigan. The first statistical analysis was used to correlate these variables with soybean yields. The data was further evaluated separately for conventional and organic land management. Previous research has shown organic management has lower yields. This is because organic production requires wider rows to enable mechanical weed control, as herbicides are not allowed for certified organic production.

The global objective of Chapter 3 was to develop a biochemical Soil Health Index (SHI). To develop this index, multivariate soybean yield prediction models were rated for their fitness based on R^2 values generated with a general and generalized linear analyses. To evaluate the strength of specific variables, a stepwise decrease of input variables was conducted in the model development. Instead of using individual variables from the multivariate model, data was put into 7 categories (land management, soil texture, environmental factors, total nitrogen, soil org. carbon, enzymes, and EL-FAME). All 7 categories represent 105 variables. Each variable was converted to a relative value constrained from 0 to 1 based on the maximum value for a given variable. To reduce the mean squared error produced by generalized linear model (GLM), a least absolute shrinkage and selection operator (Lasso) regularization step in combination with a cross-validation was performed. To automate this process the package *glmnet* was used in RStudio. This statistical analysis resulted in a

multivariate model that accounted for intercorrelations between variables and that was more robust because it was cross validated by running more than 10000 possible combinations of training vs. test data sets.

The resulting soybean yield prediction model had a R^2 value of 0.84. To develop the biochemical SHI the individual slope coefficients were separated into negative and positive groups. This information was used to calculate the weighted variable value by taking the total slope coefficient and multiplying the individual assigned weight factors with each variable data point. Because the variables and the weight factors have a range of 0 to 1, each observation resulted in a SHI score with the same range (0 to 1) after the weighted variable scores were summarized. These calculation steps were conducted separately for the positive and negative slope coefficient which included EL-FAME and enzyme coefficients. Additionally, SHI scores were determined separately for EL-FAME variables and enzyme variables which is different to the original SHI that used EL-FAME and enzyme variables.

The most common SH indicators in this study and the computed SH scores were analyzed for their ability to detect soil management at four LTES by running the Tukey's Honest Significant Difference test in combination with a sensitivity scoring algorithm. The sensitivity scores were used to identify the most sensitive SH indicators.

In Chapter 4 a total of 521 soil variables were scored for their ability to detect agricultural land management practices (e.g. crop rotation, cover cropping, soil amendments, tillage practices), restored prairies, and virgin soil in Ohio. Additionally at the agricultural scale each variable was tested for its temporal sensitivity. The most sensitive SH indicators were identified with the help of a sensitivity scoring algorithm

and their relationship to soil organic carbon was determined. The remaining SH indicators were used to determine beneficial and detrimental agricultural land management practices.

Chapter 1: Soil Health and Agricultural Challenges

1.1 Introduction

Since the beginning of the 1990s a considerable amount of research and discussions have been conducted to define and quantify SH. Between 1975 to 2023 most of the research focused on the quality of a soil reaching a total publication number of 221,337 followed by 75,436 publications related to SH (Figure 1.1). This was in part driven by the oil embargo of 1973 that impacted fertilizer availability and pricing (USDA, 2019). As such research was increased to determine whether the soil microbial community could be optimized to increase nutrient availability to crops, and in the case of nitrogen (N) to fix atmospheric N_2.

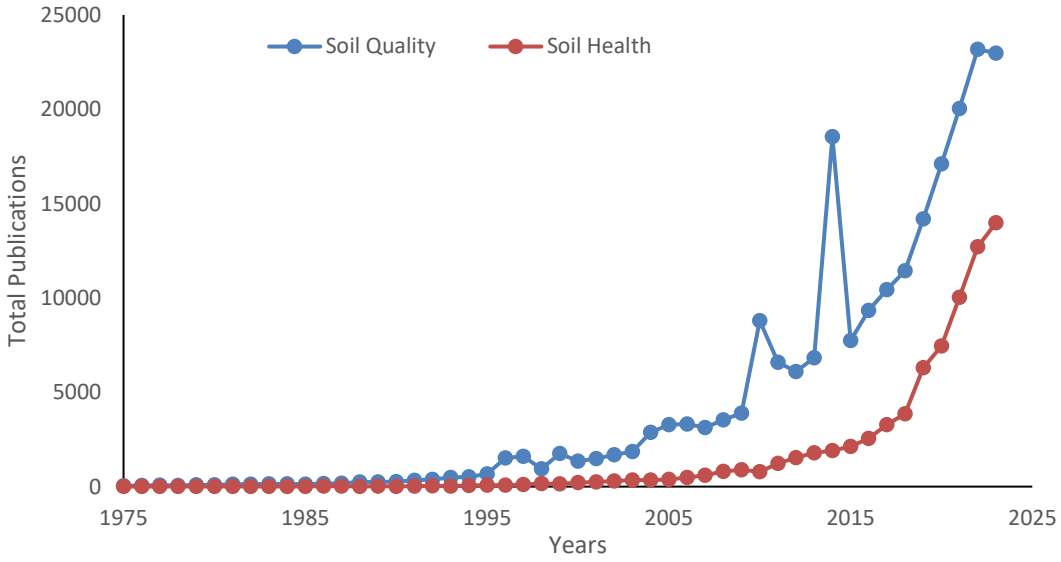


Figure 1.1 Yearly number count of publications related to the topic of "soil quality" (n=221337) and "soil health" (n=75436) from 1975-2023. Source: app.Dimensions.ai

The interest in soil quality and health is now increasing because of growing environmental challenges, interest in eliminating or reducing chemical inputs for agriculture, climate change, and the ongoing soil degradation (FAO, 2022). Another

factor is the growth of the world population which is estimated to increase by 21 % in the coming 26 years (Census.gov, Jan 2024). These challenges validate the importance of restoring and sustaining soil organic matter content, maintaining nutrient availability, and the importance of improving soil properties and processes by following the law of return (Howard, 1943, 1945; Lal, 2021).

1.2 Soil

In 2003 a case study estimated that soil degradation is negatively affecting the economy by 3-7% of the agricultural gross domestic product (AgGDP) (Berry et al., 2003). In 2021 the World Bank and OECD reported that around 4.3% of the global GDP (US$101.33 trillion) represent the AgGDP which means that the cost of soil degradation is estimated to be from 131 billion to US$305 billion per year globally. Another study estimated annual losses of US$400 billion due to the erosion of billions of tons of arable land worldwide (Borrelli et al. 2017). This means that soil degradation is causing negative losses of up to 9.2% on AgGDP. These differences in economic estimates indicate that the real costs of soil degradation are possibly higher than previously predicted.

Based on the latest global land-use numbers from 2021 for cropland (1579.9 million ha) and for pasture- and rangeland (3207.7 million ha) this breaks down to economic loss of US$162.46 per ha for cropland and US$40.01 per ha for pasture- and rangeland (FAOSTAT, 2022a). The loss of ecosystem services due to soil degradation is estimated to be even higher with losses between US$6.3 to 10.6 trillion annually which is 17 to 27 times higher than the highest reported economic losses due to soil

erosion (ELD, 2015). The economic losses represent the need for sustainable and regenerative land management practices around the globe to minimize the effect of soil erosion, degradation of soil, and loss of ecosystem services provided by soils. To achieve this goal, land management recommendations for the agricultural sector need to be based on a model that can quantify the health of the soil.

To develop such a SH model dynamic interactions between physical, chemical, and biological soil properties need to be considered (Granatstein and Bezdicek, 1992; Doran and Gregorich, 2002). Soil is a dynamic living system, that regulates ecological and environmental functions which vary across regions and soil types. Soils have diverse properties because of the unique interactions at a given site due to the major soil-forming factors, which include parent material, climate, biota, topography, and time (Buol et al., 2011). The evaluation of SH becomes more complex when climate related factors like temperature and precipitation must be considered because they can interact with soil properties (Roper et al., 2017). Ultimately SH assessment is needed to determine the impacts of land management practices such as tillage, cover cropping, manure, crop rotation, residue incorporation, and pesticides on the ability of soils to deliver agro-ecosystem services. However, detection of land management is confounded by soil type.

The earliest mention of “soil health” date back to a 1910 thesis by Henry A. Wallace who later became Secretary of Agriculture under President Franklin Roosevelt in 1933. Other works by Sir Albert Howard and Sir Robert McCarrison who wrote about the soil-human health connection and organic agriculture expanded on the concept (Brevik, 2019; Brevik and Sauer, 2015; Wallace, 1910). The work by Sir

Albert Howard introduced the Law of Return which encouraged the adoption of farming practices in which all organic waste products are recycled back to the soil and thereby would restore and sustain soil organic matter content, maintain nutrient supply, and improve soil properties and processes (Howard, 1943, 1945; Lal, 2021).A major effort to define and quantify SH started in early 1990s (Doran et al., 1994; Doran and Jones, 1996; Doran, 2002). The concept, and definition of SH is still evolving. The most recent definition of SH by the US Department of Agriculture is "The continued capacity of soil to function as a vital living ecosystem that sustains plants, animals, and humans.". The 2002 definition by John W. Doran describes SH in a more holistic way:

> *Soil health, or quality, can be broadly defined as the capacity of a living soil to function, within natural or managed ecosystem boundaries, to sustain plant an animal productivity, maintain or enhance water and air quality, and promote plant and animal health. (Doran and Gregorich, 2002)*

The terms soil quality and SH have been used in the past synonymously but the need for standardization and the desire to communicate the importance of soil to the broader public put a larger emphasis on the term SH.

Soil quality is defined as "a measure of the condition of soil relative to the requirements of one or more biological species and/or to any human purpose" (Johnson et al., 1997). In general, this means that soil quality is focused on the fitness and function of a soil to achieve a specific task. Such a task can be for example related to crop productivity, water storage or detoxification/degradation of pollutants (Doran and Zeiss, 2000).

SH on the other hand is used in a broader sense with the goal to determine what land management practices should be used to sustain not only the biological

productivity, but also sustain plant and animal health, and promote environmental quality in each ecosystem (Doran et al., 1996, 1998). This means that SH has a strong focus to identifying or quantifying the impact of sustainable practices on soils, while soil quality can be used to achieve a single ecosystem task.

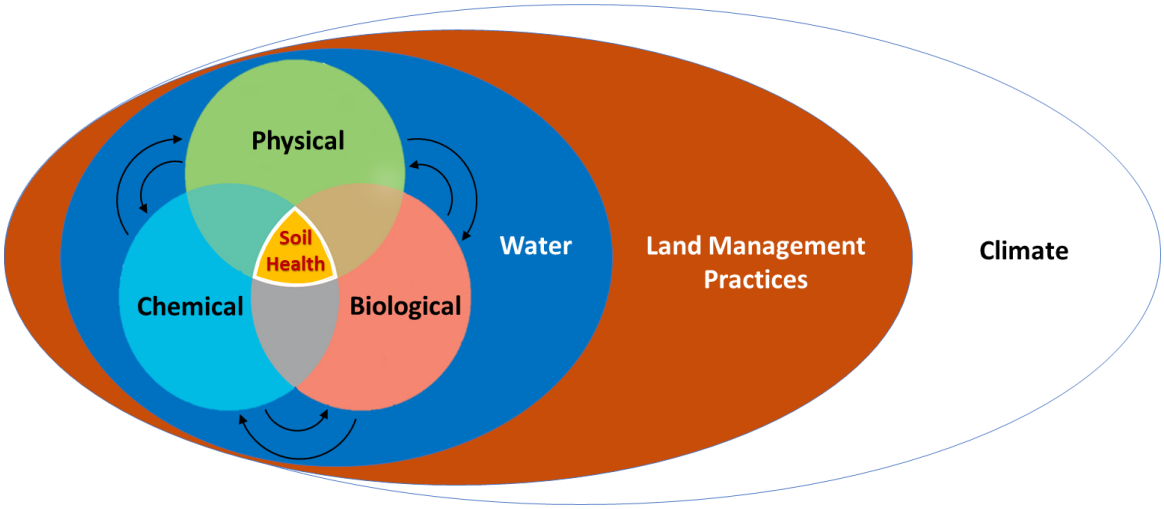


Figure 1.2 Simplified representation of the interactions between physical, chemical, and biological factors which influence and define Soil Health including the overreaching factors that can influence soil health indicators.

Over the past decades, SH initiatives, the USDA-NRCS, and the scientific community have identified and categorized soil properties that respond to specific management practices based on good, adequate, and bad thresholds (Table 1.1; Soil Health Institute, 2017; Steward et al., 2018). Furthermore, four principles and corresponding practices for improving SH have been identified and later summarized by the USDA-NRCS department: (1) maximizing continuous living roots, (2) minimizing disturbance, (3) maximizing biodiversity, and (4) maximizing soil cover (Roesch-McNally et al., 2018) (Figure 1.3).

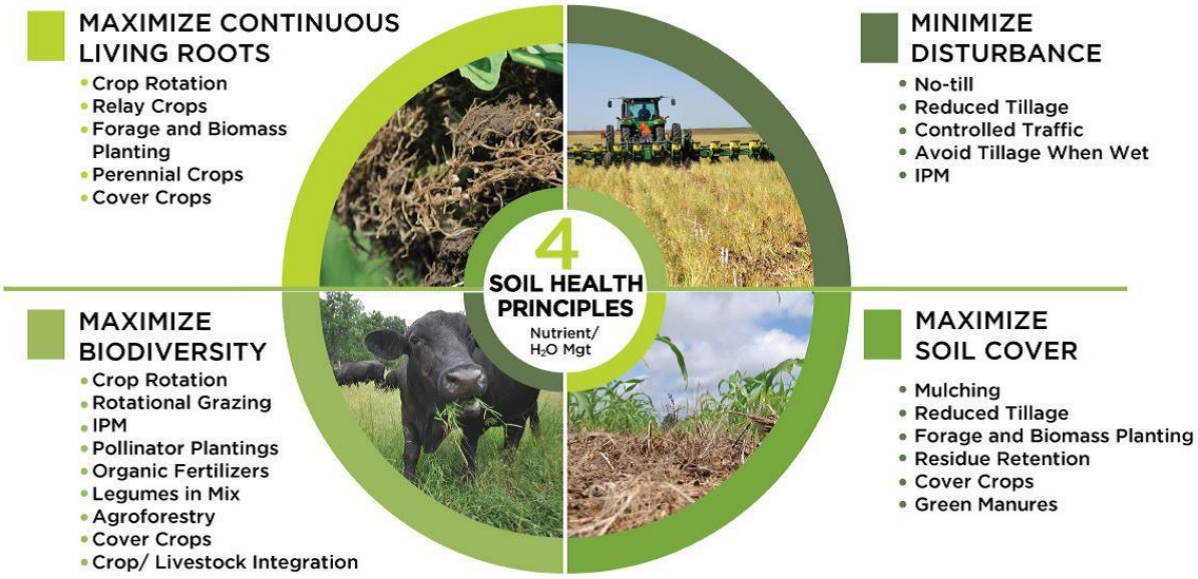


Figure 1.3 Visual representation of the four principles of Soil Health and the corresponding practices to improve the health of a soil (Roesch-McNally et al., 2018)

Based on this concept three SH assessment frameworks were developed. Instead of the common Mehlich-3 soil test, which focuses on chemical soil properties, that quantify plant available nutrients and provide a basis for fertilizer recommendations, these SH assessments further incorporate physical and biological SH indicators. The Haney Soil Health Test includes CO_2 respiration in addition to the common nutrient test (Haney et al., 2006; Haney et al., 2010). The Cornell Soil Health (CASH) test and the Soil Health Management Assessment Framework (SMAF) test use a variety of physical, chemical, and biological soil properties to evaluate the health of a soil (Andrews et al., 2004; Moebius-Clune et al., 2017). Which is shown in Table 1.2 (Norris et al., 2020). These tests provide individual indicator measurements and scoring system that integrates multiple measures into a single SH index. However, at various research sites these tests have not been able to consistently detect agricultural practices

and showed limited potential to predict crop yield (Roper et al., 2017; van Es and Karlen, 2019; Chu et al. 2019). The research sites were in North Carolina, West Tennessee, and Ontario, Canada which primarily represent humid climate regions. The outcome of these results started a discussion that the existing SH tests likely need to be calibrated on a regional basis to account for divergent soils, climates, and land management systems. (Norris et al., 2020; Ghimire et al., 2023; Chahal and Van Eerd, 2018). To find a solution for this the Soil Health Institute conducted a study on 124 diverse long-term experiment sites across North America where soil samples were collected and analyzed for 31 soil properties (Norris et al., 2020; Ghimire et al., 2023). This study concluded that three soil measurements that detected soil/crop management were: (1) soil organic carbon, (2) carbon mineralization potential, and (3) aggregate stability (Soil Health Institute, 2022). However, no scoring algorithm has been developed to quantify the health of a soil. But the evaluation based on these three soil properties is likely not sensitive or practical.

Measurable organic matter changes due to soil management typically takes decades (Jastrow, 1996), particularly in arid and semi-arid regions (Ghimire et al. 2019; Acosta-Martínez et al., 2011; Jacinthe et al. 2011; Lal, 2004). Thus, soil organic carbon changes are too slow to guide soil management.

Aggregate stability can change through extensive rain events, the formation of soil crusts, the shear strength of water flowing over the soil surface, changes in tillage practices, changes in soil amendment applications which results in soil organic carbon losses (Stavi and Lal, 2011a, 2011b; Algayer et al., 2014). It would be expected that aggregation is highly variable as a function of soil type. Also, it is labor intensive. For

these reasons aggregate stability is not a good indicator practical reasons and nor for calibration.

Wade et al. (2018) found that carbon mineralization has many sources of variation due to soil handling (sieve size), water content, and direction of rewetting. Thus, there is no standard C mineralization procedure that can be applied to enable calibration or comparison from one operator or site to another.

These shortcomings of the most prominent SH indicators shows that other indicators and model algorithms are needed that are sensitive for detecting land management, have a high throughput capability, and are seasonally stable.

To achieve this for the agricultural sector SH indicators need to be identified that have a relationship with crop yield. Past research found that healthy soils are linked to crop productivity and economic profitability (Liebig et al., 2007; Hendrickson et al., 2008; Paustian et al., 2016). This means that if agricultural land management practices follow the four principles of SH an increase in agricultural productivity should be expected and specific SH indicators should be able to reflect this (UN, 2022).

1.3 Agricultural Challenges

The identification of sensitive SH indicators is very important for the agricultural sector not only because of the challenges due to climate change, and soil degradation. But also, because the agricultural sector is slowly reaching its limits of productivity since the amount of available agricultural land peaked at 4.87 billion ha in 1999 followed by a steady decline to 4.79 billion ha by 2021 (Figure 1.4) (FOASTAT, 2022a; Ritchie and Roser, 2019; Goldewijk et al., 2017).

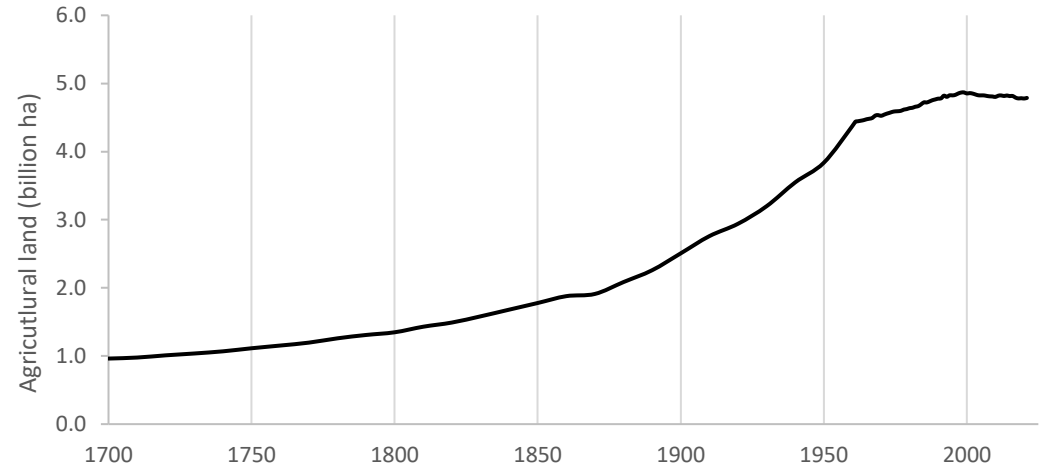


Figure 1.4 Hectares of global agricultural land from 1700 to 2021. Adapted from FOA and Goldewijk et al. (2017) data. (FAOSTAT, 2022)

This reduction of agricultural land is on average a yearly loss of 7.23 million ha of permanent meadows (1999 to 2021) and the average introduction of 3.84 million ha per year (1970-2021) of new cropland which results in a net loss of 3.39 million ha of agricultural land per year (Figure 1.5).

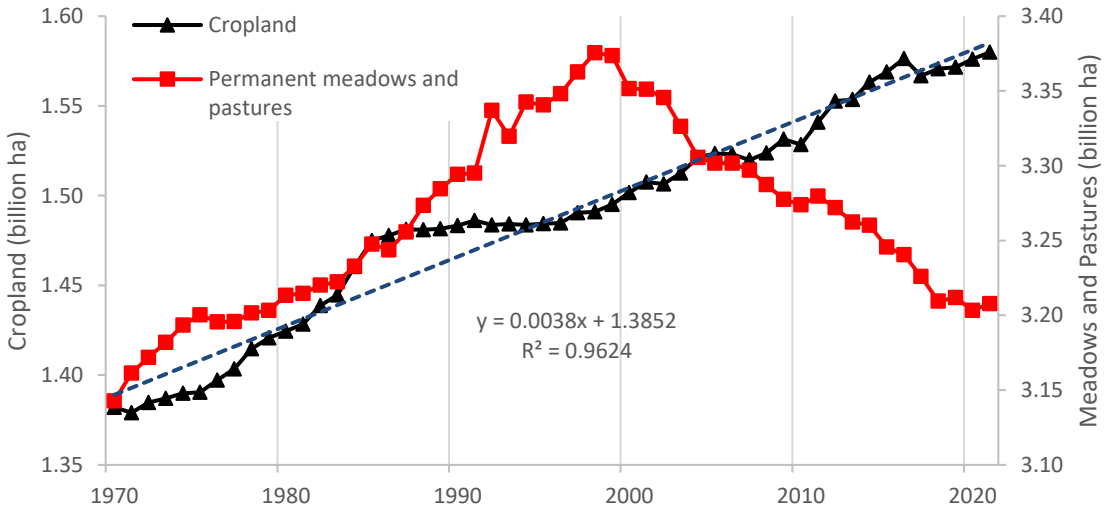


Figure 1.5 Global cropland and permanent meadows and pastures usage data between 1970 to 2021. Adapted from FOA data. (FAOSTAT, 2022)

The cropland per person calculation based on UN population data was 0.439 ha/capita of cropland in 1961. In 2021 this number dropped below 0.200 ha/capita for the first

time which represents a drop of 54%. (Figure 1.6) (FAOSTAT. 2022a; United Nations, 2022). Based on this information the agricultural output had to increase by a factor of around 2.2 in the same period. This means that the agricultural sector was able to produce not just more food but also produce it more efficiently.

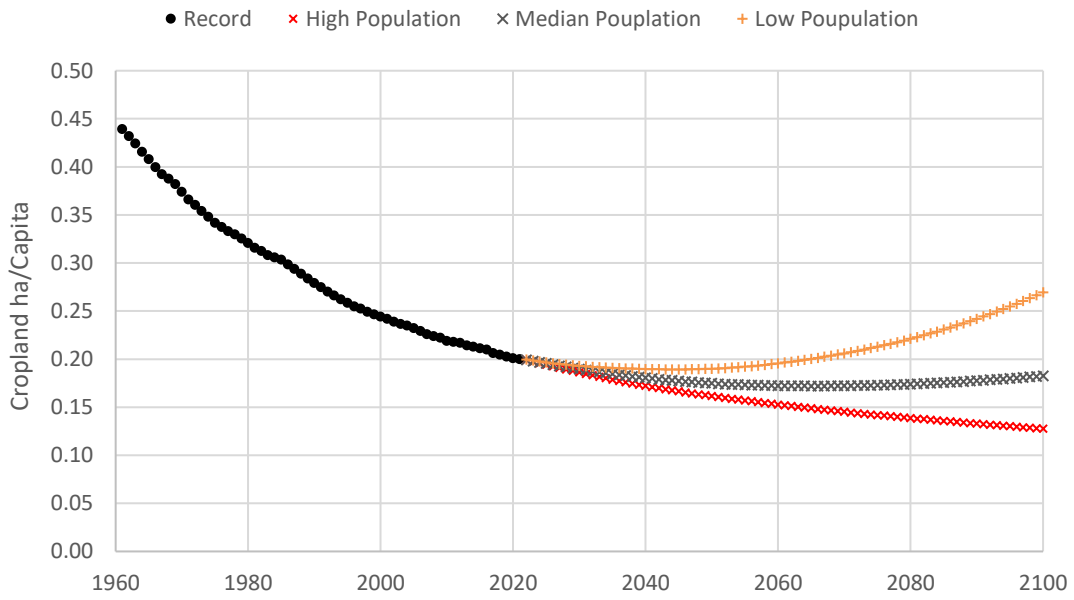


Figure 1.6 Representation of the recorded cropland per capita changes between 1960 to 2022, and the representation of three theoretical predictions model based on the assumption that cropland is going to increase linear until 2100, and population prediction number by the United Nations. Adapted from FOASTAT, 2022; UN, 2022; Census.gov, Jan 2024.

This represents a monumental achievement by farmers, scientists, and governmental organizations like the Food and Agriculture Organization of the United Nations (FAO) for combating food insecurity worldwide. Global malnourishment/malnutrition has dropped from 66% in 1950 to 24% in 1970 to 19% in 1980 to 14% in 1996 and 9% in 2020 (FAO, 1946; Boyd-Orr, 1950; FAO, 2009; Carlson, 2023; FAOSTAT, 2023). Even though the costs of a healthy diet are similar around the world today, many developing countries has major portion of the population who cannot afford these more

costly products, which in 2020 was around 3.1 billion (42.2%) people. In Africa, this is very high at 80% of the population who could not afford a healthy diet (FAO, IFAD, UNICEF, WFP and WHO, 2022). The problem of malnutrition will continue to exist with the on-going growth of the world population which is projected to reach 9.7 billion by 2050.

To achieve an increase in productivity the strain on soils will require not only an improvement in nutrient management, but also an improvement in SH for optimal productivity and food quality. As such a sensitive tool that can quantify the health of a soil is needed to provide information to farmers to guide soil management. This will require the identification of SH indicators that can: (1) respond to changes in climate and soil management quickly, (2) are easy to be sampled, measured, and interpreted by farmers, (3) are cheap and if possible accessible to farmers, and (4) can predict agronomic productivity and changes in the ecosystem services (Doran and Zeiss, 2000; Moebius-Clune et al., 2017).

Research has shown that the biological measurements for ester-linked fatty acid methyl ester (EL-FAME) biomarkers, and enzyme activities, hold potential for detecting soil/crop management effects and climate variability. (Pérez-Guzmán et al., 2021; Pérez-Guzmán et al., 2020; Bandick and Dick, 1999; Schutter and Dick, 2002; Acosta-Martínez et. al., 2014; Cotton and Acosta-Martínez, 2018; Cotton et al., 2013; Balota et al., 2004; Dick, 1984; Mbuthia et al., 2015; Li et al., 2018).

Therefore, the objectives of this dissertation were to: 1) determine the ability of EL-FAME and soil enzyme activities for detecting soil/crop practices systems in fields managed by farmers; determine if soil properties correlate with soybean yield

and develop a SH index model for detecting shifts in soil/crop management and predicting soybean yields. The hypotheses were: the selected soil biological indicators will detect soil/crop management effects and correlate with soybean yields; and that a multivariate SH indicator model can be developed that can predict soybean yields at a high level of probability which then can be used to develop a sensitive biochemical SH indicator.

1.4 References

Acosta-Martínez, V., Lascano, R.; Calderón, F., Booker, J.D., Zobeck, T.M., Upchurch, D.R., 2011. Dryland cropping systems influence the microbial biomass and enzyme activities in a semi-arid sandy soil. Biol. Fertil. Soils, 47, 655–667.

Acosta-Martínez, V., Moore-Kucera, J., Cotton, J., Gardner, T., & Wester, D., 2014. Soil enzyme activities during the 2011 Texas record drought/heat wave and implications to biogeochemical cycling and organic matter dynamics. Applied Soil Ecology, 75, 43–51. http://doi.org/10.1016/j.apsoil.2013.10.008

Algayer, B., Wang, B., Bourennane, H., Zheng, F., Duval, O., Li, G., Le Bissonnais, Y., Darboux, F., 2014. Aggregate stability of a crusted soil: Differences between crust and sub-crust material, and consequences for interrill erodibility assessment: An example from the Loess Plateau of China. Eur. J. Soil Sci. 65, 325–335. https://doi.org/10.1111/ejss.12134

Andrews, S. S., Karlen, D. K., & Cambardella, C. A., 2004. The soil management assessment framework: A quantitative soil quality evaluation method. Soil Science Society of America Journal, 68, 1945–1962.

Bandick, A. K., & Dick, R. P., 1999. Field management effects on soil enzyme activities. Soil Biology & Biochemistry, 31, 1471–1479.

Balota, E. L., Kanashiro, M., Colozzi Filho, A., Andrade, D. S., & Dick, R. P., 2004. Soil enzyme activities under long-term tillage and crop rotation systems in subtropical agro-ecosystems. Brazilian Journal of Microbiology, 35, 300–306 http://doi.org/10.1590/S1517-83822004000300006

Berry, L., Olson j. and Campbell D., 2003. Assessing the Extent, Cost and Impact of Land Degradation at the National Level: Findings and Lessons Learned from Seven Pilot Case Studies, commissioned by Global Mechanism with support from the World Bank.

Borrelli, P., Robinson, D.A., Fleischer, L.R. et al., 2017. An assessment of the global impact of 21st century land use change on soil erosion. Nat Commun 8, 2013 https://doi.org/10.1038/s41467-017-02142-7

Boyd-Orr, L. J., 1950. The Food Problem. Scientific American, 183(2), 11–15. http://www.jstor.org/stable/24967518

Brevik, E.C., Steffan J.J., Burgess L.C., and Cerdà A., 2017. Links between soil security and the influence of soil on human health, in: D. Field, C. Morgan, and A. McBratney (Eds.), Global Soil Security. Progress in Soil Science Series, Springer, Rotterdam. p. 261-274.

Brevik, E., 2019. A Brief History of the Soil Health Concept.

Buol, S.W., R.J. Southard, R.C. Graham, and P.A. McDaniel., 2011. Soil genesis and classification. 6th ed. Wiley-Blackwell, Oxford, UK.

Carlson A., 2023. Hunger Is Not Eradicated: The Food Crisis in Africa. Origins Current Events in Historical Perspective, https://origins.osu.edu/read/hunger-not-eradicated-food-crisis-africa?language_content_entity=en

Chahal, I., & Van Eerd, L. L. (2019). Quantifying soil quality in a horticultural-cover cropping system. *Geoderma*, *352*, 38–48. https://doi.org/10.1016/j.geoderma.2019.05.039

Chu, M., Singh, S., Walker, F. R., Eash, N. S., Buschermohle, M. J., Duncan, L. A., & Jagadamma, S. (2019). Soil Health and Soil Fertility Assessment by the Haney Soil Health Test in an Agricultural Soil in West Tennessee. Communications in Soil Science and Plant Analysis, 50(9), 1123–1131. https://doi.org/10.1080/00103624.2019.1604731

Cotton, J., & Acosta-Martínez, V., 2018. Intensive tillage converting grassland to cropland immediately reduces soil microbial community size and organic carbon. Agricultural & Environmental Letters, 3, 1– 4.

Cotton, J., Acosta-Martínez, V., Moore-Kucera, J., & Burow, G., 2013. Early changes due to sorghum biofuel cropping systems in soil microbial communities and metabolic functioning. Biology and Fertility of Soils, 49, 403–413. http://doi.org/10.1007/s00374-012-0732-z

Dick, W. A., 1984. Influence of long-term tillage and crop rotation combinations on soil enzyme activities. Soil Science Society of America Journal, 48, 569–574. http://doi.org/10.2136/sssaj1984.03615995004800030020x

Doran, J.W., Sarrantonio, M., Janke, R., 1994. Strategies to promote soil quality and soil health. In: Pankhurst, et al. (Eds.), Soil Biota: Management in Sustainable Farming Systems. CSIRO, Vic., Australia, pp. 230–237.

Doran, J.W. and Jones, A.J., 1996. Methods for Assessing Soil Quality. SSSA Publication No. 49, Soil Science Society of America, Madison, WI. - References - Scientific Research Publishing. (n.d.). Www.scirp.org. Retrieved July 16, 2024, from https://www.scirp.org/reference/referencespapers?referenceid=2023241

Doran, J.W., Sarrantonio, M., Liebig, M., 1996. Soil health and sustainability. In: Sparks, D.L. (Ed.), Advances in Agronomy, Vol. 56, Academic Press, San Diego, CA, USA, pp. 1–54.

Doran, J.W., Liebig, M.A., Santana, D.P., 1998. Soil health and global sustainability. In: Proceedings of the 16th World Congress of Soil Science. Montepellier, France, 20-26 August 1998.

Doran, J. W., & Zeiss, M. R., 2000. Soil health and sustainability: managing the biotic component of soil quality. Applied Soil Ecology, 15(1), 3–11. https://doi.org/10.1016/s0929-1393(00)00067-6

Doran, J.W., Gregorich, E.G., 2002. Soil quality and sustainable agriculture. In: Lal, R. (Ed.), Encyclopedia of Soil Science. Marcel Dekker, New York, NY, in Press.

ELD Initiative, 2015. The value of land: Prosperous lands and positive rewards through sustainable land management. Available from www.eld-initiative.org.

FAO (Food and Agriculture Organization of the United Nations). 1946. World Food Survey. Rome, https://www.fao.org/publications/card/en/c/cb6105en/

FAO (Food and Agriculture Organization of the United Nations). 2009. The State of the pf Food Insecurity in the World. Main report. Rome, https://www.fao.org/3/i0876e/i0876e00.htm

FAO (Food and Agriculture Organization of the United Nations). 2022. The State of the World's Land and Water Resources for Food and Agriculture: Systems at Breaking Point. Main report. Rome, https://www.fao.org/documents/card/en?details=cb9910en

FAO, IFAD, UNICEF, WFP and WHO. 2022. The State of Food Security and Nutrition in the World 2022. Repurposing food and agricultural policies to make healthy diets more affordable. Rome, FAO. https://doi.org/10.4060/cc0639en

FAOSTAT. 2022. Land Use. In: FAO. Rome. https://www.fao.org/faostat/en/#data/RL

FAOSTAT. 2023. Suite of Food Security Indicators. In: FAO. Rome www.fao.org/faostat/en/#data/FS

Ghimire, R., Ghimire, B., Mesbah, A.O., Sainju, U.M., Idowu, O.J., 2019. Soil health response of cover crops in winter wheat–fallow system. Agron. J., 111, 2108–2115.

Ghimire R, Thapa VR, Acosta-Martinez V, Schipanski M, Slaughter LC, Fonte SJ, Shukla MK, Bista P, Angadi SV, Mikha MM, et al., 2023. Soil Health Assessment and Management Framework for Water-Limited Environments: Examples from the Great Plains of the USA. Soil Systems. ; 7(1):22. https://doi.org/10.3390/soilsystems7010022

Goldewijk K., Beusen K., A., Doelman J., Stehfest E., 2017, Anthropogenic land use estimates for the Holocene; HYDE 3.2, Earth System Science Data, 9, 927-953.

Granatstein, D., and Bezdicek D. F., 1992. The need for a soil quality index: local and regional perspectives. Am. J. Altem. Agric. 7:12-16.

Haney, R. L., Haney, E. B., Hossner, L. R., & Arnold, J. G., 2006. Development of a New Soil Extractant for Simultaneous Phosphorus, Ammonium, and Nitrate Analysis. Communications in Soil Science and Plant Analysis, 37(11-12), 1511–1523. https://doi.org/10.1080/00103620600709977

Haney, R., Haney, E., Hossner, L., & Arnold, J., 2010. Modifications to the new soil extractant H3A-1: A multinutrient extractant. Communications in Soil Science and Plant Analysis, 41(12), 1513–1523. https://doi.org/10.1080/00103624.2010.482173.

Hendrickson, J.R.; Liebig, M.A.; Sassenrath, G.F., 2008. Environment and integrated agricultural systems. Renew. Agric. Food Syst., 23, 304–313.

Howard, A., 1943. An Agricultural Testament. New York: Oxford University Press.

Howard, A., 1945. The Soil and Health: A Study of Organic Agriculture. Lexington, KY: University Press of Kentucky.

Jacinthe, P.-A., Shukla, M.K., Ikemura, Y., 2011. Carbon pools and soil biochemical properties in manure-based organic farming systems of semi-arid New Mexico. Soil Use Manag, 27, 453–463.

Johnson, D.L., Ambrose, S.H., Bassett, T.J., Bowen, M.L., Crummey, D.E., Isaacson, J.S., Johnson, D.N., Lamb, P., Saul, M., Winter-Nelson, A.E., 1997. Meanings of environmental terms. J. Environ. Qual. 26, 581–589.

Lal, R., 2020. The Soil-Human Health-Nexus (1st ed.). CRC Press. https://doi.org/10.1201/9780367822736

Lal, R., 2004. Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. Science, 304(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Li, C., Fultz, L. M., Moore-Kucera, J., Acosta-Martínez, V., Kakarla, M., & Weindorf, D. C., 2018. Soil microbial community restoration in Conservation Reserve Program semi-arid grasslands. Soil Biology & Biochemistry, 118, 166–177.

Liebig, M.A.; Tanaka, D.L.; Krupinsky, J.M.; Merrill, S.D.; Hanson, J.D. Dynamic cropping systems: Contributions to improve agroecosystem sustainability., 2007. Agron. J., 99, 899–903.

Mbuthia, L. W., Acosta-Martínez, V., DeBruyn, J., Schaeffer, S., Tyler, D., Odoi, E., Mpheshea, M.,Walker, F.,&Eash, N., 2015. Long term tillage, cover crop, and fertilization effects on microbial community structure, activity: Implications for soil quality. Soil Biology & Biochemistry, 89, 24–34.

Moebius-Clune, B., Moebius-Clune, D., Gugino, B., Idowu, O., Schindelbeck, R., Ristow, A. J., Abawi, G. S., 2017. Comprehensive assessment of soil health, the Cornell framework (3rd ed.) Ithaca, NY: Cornell University.

Norris CE, Bean GM, Cappellazzi SB, et al., 2020. Introducing the North American project to evaluate soil health measurements. Agronomy Journal ; 112: 3195–3215. https://doi.org/10.1002/agj2.20234

UN, World Population Prospects., 2022 – processed by Our World in Data. “Population” [dataset]. United Nations, World Population Prospects (2022) https://ourworldindata.org/population-growth

USDA. 2019. Recommended Soil Health Indicators and Associated Laboratory Procedures. Technical Note No. 450-03.

Paustian, K.; Lehmann, J.; Ogle, S.; Reay, D.; Robertson, G.P.; Smith, P., 2016. Climate-smart soils. Nature, 532, 49–57.

Pérez-Guzmán L, Phillips LA, Acevedo MA, Acosta-Martínez V., 2021 Comparing biological methods for soil health assessments: EL-FAME, enzyme activities, and qPCR. Soil Sci Soc Am J.; 85: 636–653. https://doi.org/10.1002/saj2.20211

Pérez-Guzmán, L., Acosta-Martínez, V., Phillips, L. A., & Mauget, S. A. 2020. Resilience of the microbial communities and their enzymatic activities in semi-arid agricultural soils during natural climatic variability events. Applied Soil Ecology, 149. https://doi.org/10.1016/j.apsoil.2019.103487

Ritchie H., Roser M., 2019. Land Use, Published online at OurWorldInData.org. https://ourworldindata.org/land-use

Roesch-McNally G., Moore-Kucera J., Owen C., 2018. Applying Soil Health Management Systems to Reduce Climate and Weather Risks in the Northwest, USDA-NRCS, https://www.climatehubs.usda.gov/sites/default/files/Mngmt%20Systems%20Reduce%20Risks%20in%20the%20NW.pdf

Roper, W.R., Osmond, D.L., Heitman, J.L., Wagger, M.G. and Reberg-Horton, S.C., 2017. Soil Health Indicators Do Not Differentiate among Agronomic Management Systems in North Carolina Soils. Soil Science Society of America Journal, 81: 828-843. https://doi.org/10.2136/sssaj2016.12.0400

Schutter, M. E., & Dick, R. P., 2002. Microbial community profiles and activities among aggregates of winter fallow and cover-cropped. Soil Science Society of America Journal, 66, 142–153. http://doi.org/10.2136/sssaj2002.1420

Stavi, I., Lal, R., 2011a. Variability of soil physical quality and erodibility in a watereroded cropland. Catena 84, 148–155. https://doi.org/10.1016/j.catena.2010.10.006

Stavi, I., Lal, R., 2011b. Loss of soil resources from water-eroded versus uneroded cropland sites under simulated rainfall. Soil Use Manag. 27, 69–76. https://doi.org/10.1111/j.1475-2743.2010.00312.x

Stewart, R. D., Jian, J., Gyawali, A. J., Thomason,W. E., Badgley, B. D., Reiter, M. S., & Strickland, M. S., 2018. What we talk about when we talk about soil health. Agricultural & Environmental Letters, 3, 180033. https://doi.org/10.2134/ael2018.06.0033

Soil Health Institute, 2017. Enriching Soil, Enhancing Life an Action plan for Soil Health https://soilhealthinstitute.org/app/uploads/2021/10/Action-Plan-FINAL-for-flipbook-3.pdf

Soil Health Institute, 2018a. The North American Project to Evaluate Soil Health Measurements (NAPESHM) Tier 1 list. https://soilhealthinstitute.org/?attachment_id=2547&highlight=napeshm&_rt=MnwxfG5hcGVzaG18MTcwNDk5MDM0MQ&_rt_nonce=fb3ff46d2b

Soil Health Institute, 2018b. The North American Project to Evaluate Soil Health Measurements (NAPESHM) Tier 2 list. https://soilhealthinstitute.org/?attachment_id=2548&highlight=napeshm&_rt=M3wxfG5hcGVzaG18MTcwNDk5MDM0MQ&_rt_nonce=c52dbbeed7

Soil Health Institute, 2022. Recommended Measurements for Scaling Soil Health Assessment https://soilhealthinstitute.org/app/uploads/2022/10/SHI_SoilHealthMeasurements_factsheet.pdf.

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), 721–732. https://doi.org/10.2136/sssaj2018.09.0338

Wade, J., Culman, S. W., Hurisso, T. T., Miller, R. O., Baker, L., & Horwath, W. R. (2018). Sources of Variability that Compromise Mineralizable Carbon as a Soil Health Indicator. Soil Science Society of America Journal, 82(1), 243–252. https://doi.org/10.2136/sssaj2017.03.0105

Wallace, H.A. 1910. Relation between livestock farming and the fertility of the land. Unpublished thesis. Iowa State University, Ames, IA.

Table 1.1 Soil Health indicators adapted from the Soil Health Institute tier list. Each indicator is organized into the corresponding biological, physical, chemical, and other measurement standard (Soil Health Institute 2018a, 2018b).

Standard	Tier 1 - Indicator		Tier 2 - Indicator	Tier 3 - Indicator
Biological	▪ Total Nitrogen ▪ Carbon Mineralization ▪ Nitrogen Mineralization	*Soil Org. Carbon / Soil Org. Matter* ***(Core of Soil Health)***	▪ Enzymes: β-glucosidase N-acetyl-B-D-glucosiminidase Phosphomonoesterase Arylsulfatase ▪ Phospholipid Fatty Acid (PLFA) ▪ Soil Protein Index ▪ Active Carbon ▪ Genomics	An indicator that has potential to add significant information about soil health in specific locations or on large scales, but specific relationships among measured values, soil processes, and effects of land management are not fully understood. (Soil Health Institute, 2017)
Physical	○ Soil Texture ○ Bulk density ○ Infiltration Rate ○ Available Water Holding Capacity ○ Water-Stable Aggregation ○ Penetration Resistance ○ Soil Electrical Conductivity		○ Aggregate Stability ○ Soil Stability Index	
Chemical	▪ Soil pH ▪ Cation Exchange Capacity ▪ Base Saturation ▪ Extractable Phosphorus ▪ Extractable Potassium ▪ Micronutrients		▪ Sodium Adsorption Ratio	
Other	▪ Crop Yield ▪ Erosion Rating		▪ Reflectance (Spectroscopy)	

Table 1.2 List of soil property measurements that are used in three individual Soil Health tests. Adapted from Norris et al. (2020).

Soil Health Test	Measurement
Soil Management Assessment Framework	• Nematode Maturity Index • Metabolic Quotient Determined from Soil Respiration and Microbial Biomass • Bulk Density • Total organic Carbon • Microbial biomass Carbon • Potentially mineralizable Nitrogen • Soil pH • Soil Test P • Macroaggregate Stability • Soil Depth • Available Water Holding Capacity • Electrical Conductivity • Sodium Adsorption Ratio
Comprehensive Assessment of Soil Health	• Soil Texture - modified method utilizing sieves and decanting • Available Water Holding Capacity • Surface Hardness by Penetrometer • Subsurface Hardness by Penetrometer • Aggregate Stability by Rainfall Simulator • Soil Organic Matter by Loss on Ignition • Soil Protein by Autoclaved Citrate Extractable Protein Index • Soil Respiration by CO2–C analysis following 4-d incubation of moist soil • Active C by Colorimetric Changes to K Permanganate Solution • Soil pH by 1:2 soil water suspension • Basic Extractable P, K, Mg, Fe, Zn • Enhanced Extractable Al, As, B, Ba, Ca, Cd, Co, Cr, Cu, Mn, Na, Ni, Pb, S, Sr by modified Morgan's Solution (Ammonium Acetate and Acetic Acid)
The Soil Health Tool (Haney Test)	• CO2–C Analysis following 24-h Incubation of Moist Soil • Water Extractable organic Carbon and Nitrogen • Oxalic, malic, and Citric Acid (H3A) extractable P, K, Mg, Ca, Na, Zn, Fe, Mn, Cu, S, and Al • Total water and H3A extractable NO3–N, NH4–N, and PO4–P

Chapter 2: Analysis of Individual Soil Health Indicators to predict Soybean Yield

2.1 Abstract

Soil Health (SH) indicators are needed to guide and quantify sustainable soil management systems. Furthermore, the health of soil should correspond to crop yield. Although certain soil properties have been shown to detect crop/soil management, relating these to yield has been minimally investigated and the most part unsuccessful. Certain enzyme assays and the ester-linked fatty acid methyl ester (EL-FAME) microbial biomarkers have been shown to be sensitive indicators of soil/crop management but are largely uninvestigated for their potential to predict crop yield. Therefore, the objective was to conduct linear regression analysis to determine their relationship to soybean yield, a major crop world-wide. These along with total carbon, total nitrogen, soil organic carbon, pH and texture were measured on soils collected annually over three years at 106 farm sites in Ohio, and for one year at four long-term experimental sites (n=47). The biological indicators were normalized for clay or sand content to overcome soil texture variability as a confounding factor for calibrating these, independent of soil type. Additionally, environmental factors (e.g. precipitation, soybean growth time) and land management variables (e.g. quantitative tillage practices, quantitative residue coverage scores) were regressed against soybean yield. Independent variables were also transformed to account for skewed distribution curves and reevaluated. Linear regression models were developed separately for the conventional and organic data sets. For the complete data set (n=153) the analysis showed the highest R^2 of 0.36 for the categorical variable that differentiated between

organic and conventional farming practices. The next nine variables were identified to have fair coefficients of determination. For the organic data set (n=30) the regression analysis showed that the EL-FAME biomarker 17:0 (R^2 = 0.29), 17:0 10-ME (R^2 = 0.29), cy19:0 (R^2 = 0.28), and soil organic carbon (SOC) (R^2 = 0.26) had a moderate coefficient of determination. For the conventional data set (n=123) the ten highest variables resulted in poor coefficients of determination with total nitrogen having the highest R^2 of 0.08. The regression analysis showed that agronomic productivity predictions need to be done separately for organic and conventional data sets. Most soil properties had fair or poor relationship with soybean yield which did not improve by normalizing these properties with texture (dividing property by sand or clay content). Overall, this study showed that soybean yields cannot be predicted by one independent variable when a diverse data set is used.

2.2 Introduction

Soybeans (Glycine max) are a major crop in the US (33.6 million ha) and grown on soils that are being eroded that have lost 50 to 66% of soil organic carbon (Lal, 2004a). To remediate soil sustainable management practices such as no-till, cover cropping, and organic amendments need to be implemented which improve the health of a soil. Healthy soils are expected to be productive, maintain environmental quality, and promote plant and animal health (Doran and Parkin, 1994). To achieve this various soil health indicators have been proposed that would allow us to quantify soil health.

This includes the Haney Soil Health test and the Comprehensive Assessment of Soil Health test (CASH; Cornell Soil Health Laboratory) (Haney et al., 2006;

Moebius et al., 2007). Both tests analyze a set of biological, chemical, and physical soil properties (Norris et al., 2020). Compared to the traditional Mechlich-3 soil test that analyzes mineral nutrients to provide plant nutrient recommendations, the intention of these soil health tests is a holistic assessment of a soil to reflect its ability to deliver ecosystem services. However, a study conducted in 2017 in North Carolina found that the Haney and CASH tests had limited ability for detecting crop/soil management effects, and that they were not correlated with crop yields (Roper et al., 2017). Studies on these indicators conducted in West Tennessee, Pennsylvania, and in Ontario, Canada came to the same conclusion (Chu et al., 2019; Chahal and Eerd, 2018; Faé et al., 2020).

The Soil Health Institute study collected soil samples from 124 long-term experiment sites (LTES) from Canada to Mexico where each site had various soil/crop management treatments (Norris et al., 2020). This study recommended that soil organic carbon (SOC), carbon mineralization potential, and aggregate stability to soil health indicators (Bagnall et al, 2023; Liptzin et al. 2022; Rieke et al., 2022). However, there was no relationship of crop yields with any of these indicators reported. Previous publications suggested that it is unlikely that a single or a set of indicators would correlate with crop yields because of the wide variation in soil types, climates, environmental conditions, and agricultural land management practices (Liebig et al., 2001; Lehman et al., 2015; Ghimire et al., 2023).

Certain soil biological properties have been shown to be good SH indicators/ This includes the activities of three hydrolytic enzyme assays that have been found to be sensitive for detecting soil management in periods as short as 1- 3 years.

Furthermore, they play key roles in the C (β-glucosidase; EC 3.2.1.21 b-D-glucoside glucohydrolase), C and N (β-glucosaminidase; EC 3.2.1.52 *p*-nitrophenyl-N-acetyl-β-D-glucosaminide), and S (arylsulfatase; EC 3.1.6.1 arylsulfate sulfohydrolase) cycles. They have detected the following management or environmental conditions: cover cropping, organic amendments, heavy metals, herbicides, tillage, and/or perennial management (see review by Acosta-Martínez, 2021; Ghimire et al. 2023). For arylsulfatase and β-glucosidase, Bandick and Dick (1999) and Ndiaye et al. (2000) have shown in-season stability which is an asset over other microbial properties that have wide variability due to environmental factors that would be difficult to calibrate.

Another potential microbial SH indicator is the Ester Linked - Fatty Acid Methyl Ester (EL-FAME) analysis. It has been widely used to profile microbial community composition because it is cost effective, less time-consuming, and more easily interpreted/quantified than DNA-based methods. The EL-FAME method measures biomarkers that are structural components of all microbial cell membranes that represent broad functional microbial groups (Zelles et al., 1994). This method extracts fatty acids from soil samples and converts them to FAMEs using an alkaline reagent (Schutter & Dick, 2000). This analysis produces representative biomarkers for Gram positive and negative bacteria, Actinobacteria, saprophytic fungi, arbuscular mycorrhizal fungi (AMF), and eukaryotes (Zelles, 1997, 1999). Fatty acid profiling methods are thought to represent the "active" microbial biomass (Pinkart et al., 2002; White et al., 1996) and have been shown to reflect early changes in microbial community structure due to management practices (Bhandari et al., 2018; Schutter et al., 2001; Schutter, and Dick, 2002; Bossio & Scow, 1998; Zhang et al., 2016).

However, for these enzyme assays and EL-FAME analysis there is very little information available from on-farm research and on how well they correlate with crop yields.

Therefore, the objective of this study was to: (1) determine if a single soil property, land management, or environmental variable can predict soybean yields, and (2) assess if an indicator would improve its coefficient of determination (R^2) after the data set is separated into conventional and organic observation fragments. To test this, soil samples were collected at conventional and certified organic on-farm sites located in eight counties of Ohio, and at four long-term field experimental sites (LTES). The hypothesis was that there would be a strong predictive relationship of EL-FAME biomarkers or enzyme activities with soybean yield, evidenced by high coefficient of determinations (R^2 = > 0.50).

2.3 Material and Methods

2.3.1 Study Sites

A total of 153 soil samples were collected each spring over the period from 2019 to 2021. Each sample site was identified, and GPS tracked before any sampling occurred. One hundred and six samples came from 18 farms in eight counties in Ohio (Clinton, Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas). In 2021 47 soil samples were collected from four LTES. Three of the LTES are in Ohio and one in Michigan. The soils are classified as: silt loam (59%), as a loam (20%), as a clay loam (11%), and as a silty clay loam (10%).

In most cases there were at least two fields where one field had soybeans

(Glycine max) the first year and the second field had soybeans the following year of sampling. Each farmer was surveyed in person or over the phone to get information on past land management on each field and management plans for each growing season. Five farmers are certified organic farmers and the remaining 13 have conventional land management.

2.3.2 Organic Farm Sites

The organic farm sites (n=18) were in Madison, Handcock, and Clinton county and have been under agricultural management for 50 to more than 100 years. Fields under organic management had been in place from to one to 20 years and range in size from 5 to 47 ha. Precipitation for each site ranged from 244 to 553 mm (Climate Fieldview, n.d). The growing period from planting to harvest ranged from 120 to 160 days. Most organically managed soils were a Crosby-Lewisburg silt loams (mesic Aeric Epiaqualfs / shallow Aquic Hapludalfs) and four were Mollisols (Soil Survey Staff, 2019). Furthermore, the only four Mollisols in the study were identified at two separate organic farm field locations. All organic farm sites used organic seeds, had a soybean-corn-wheat rotation, a 30-inch (76 cm) row spacing, and no synthetic inputs to meet certified organic standards. However, across the organically, managed fields there was variation in tillage manure applications, and cover cropping.

2.3.3 Conventional Farm Sites

Thirteen conventionally managed fields ranged in size from 4 to 77 ha. Sixteen fields had a soybean-corn-wheat rotation and 72 a soy-corn rotation.

Conventional farmers had a 15-inch (38 cm) (n=74) or 7.5-inch (19 cm) (n=12) row spacing. All conventional farmers used synthetic fertilizers and herbicides. Herbicide management was divided in three application categories: only glyphosate (*N-(phosphonomethyl) glycine*), glyphosate with a secondary herbicide, and dicamba (*3,6-dichloro-2-methoxybenzoic acid).* All conventionally managed fields had Alfisols (Soil Survey Staff, 2019). The soil samples came from Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas counties. The agricultural fields under conventional management ranged from 1 to 100 years of usage. One field was converted from native land to farmland. Seasonal precipitation between planting and harvest was recorded using the Climate Fieldview website which ranged from 329 to 653 mm. (Climate Fieldview, n.d). The individual recorded precipitation levels for conventional farm sites varied between 329 to 653 mm. The growth period ranged from 114 to 173 days. Other practices varied for cover cropping, manure application rates and type, and tillage.

2.3.4 Long Term Field Sites

2.3.4.1 Wooster - Triplett-Van Doren Site

The LTES in Wooster, OH (40.764° N, -81.906° W) was established in 1962 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Wooster silt loam (fine-loamy, mixed, active, mesic Oxyaquic Fragiudalfs) with a 2-6 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25-30 % for sand, 55-60 % for silt and 15% for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). Deiss et al. (2021) reported a range of 5.4 to 6.8 for soil

pH.

The experimental has a two-way factorial randomized complete block design with three replications with three tillage treatments, and three crop rotations (Dick and Van Doren Jr., 1985; Deiss et al., 2021). Plot size is 22.3 m by 4.3 m.

The three tillage treatments are: (1) no-tillage (NT); (2) chisel (minimum) tillage (CT); or (3) moldboard plow (MP). The minimum tillage treatment had a para plow from 1962 to 1984, after which a chisel cultivator was used. Chisel tillage loosens the soil and allows up to 30% litter retention on the soil surface. Moldboard tillage inverts soil to a depth of 20 cm and buries the litter, leaving 5 % or less on the soil surface (Dick et al., 2013).

The three crop rotation treatments on the site are: (1) continuous corn *(Zea mays L.)* (CC); (2) corn and soybean *(Glycine max L.)* (CS); and (3) corn and oat *(Avena sativa L.)* and/or alfalfa *(Medicago sativa)* or clover *(Trifolium repens L.)* (CFF). Nine soil samples came from the CS rotation plots which were collected in 2021.

2.3.4.2 Hoytville - Triplett-Van Doren Site

The LTES in Hoytville, OH (41.222 ° N, -83.762° W) was established in 1963 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Hoytville clay loam (fine, illitic, mesic Mollic Epiaqualfs) with a 0-1 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25 % for sand, 39 % for silt and 36 % for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). In contrast to the Wooster soil, The Hoytville soil has a poor surface and internal drainage, and it cracks when dry. In 1952 a subsurface tile drainage was installed at a

depth of 1.2 - 1.4 m (Dick et al., 1986b; Deiss et al., 2021). Deiss et al. (2021) reported a range of 4.3 to 7.5 for soil pH.

It has a two-way factorial randomized complete block design with three replications, and the identical three tillage treatments, and three crop rotations as the Wooster LTES (Dick and Van Doren Jr., 1985; Deiss et al., 2021). The plot size is 30.5 m by 6.4 m. Eight soil samples came from the CS rotation plots which were collected in 2021.

2.3.4.3 Columbus - Straw Mulch Experiment

The Straw Mulch Experiment (40.017° N, -83.0395° W) was established in 1996 by the Carbon Management and Sequestration Center (CMASC) at the Ohio State University. The objective of this LTES is to determine the effect of wheat straw *(Triticum aestivum L.)* mulching on soil quality, soil organic carbon (SOC) sequestration and dynamics, and greenhouse gas emissions (Blanco-Canqui and Lal, 2007). No mechanical tillage is used, and glyphosate *(N-Phosphonomethyl glycine)* is used to control weeds. The primary soil series is a Crosby silt loam (fine, mixed, active, mesic Aeric Epiaqualfs) with a 2-6 % slope (Soil Survey Staff, 2019). For the top 15 cm the soil particle size is 22-23 % for sand, 53-56 % for silt, and 22-24 % for clay (Soil Survey Staff, 2019; Nawaz et al., 2016; Saroa and Lal, 2003). Measured soil pH at a depth of 0 to 15 cm ranged from 5.7 to 7.1.

The experimental design is a two-way factorial completely randomized block design (3 replications) with three mulch rates and two fertilizer rates. The fry mulch treatments are: (1) no mulch (control), (2) 8 Mg ha^{-1} yr^{-1}, and (3) 16 Mg ha^{-1} yr^{-1}. The

fertilizer treatments are: (1) no fertilizer application (control), or (2) annual broadcast fertilizer application with a rate of 244 kg N ha^{-1} (184 kg N ha^{-1} as Urea) and 60 kg ha^{-1} of NPK). Each year, the wheat straw is applied in the spring followed by fertilizer application in the late spring to early summer. Until 2020 no crops were grown on the plots after which for two years corn and soybean were grown on them. Plot size is 5 by 5 m. Each plot on which the crop experiment took place was separated into two halves (2.5 by 5 m) with a corn-corn and soybean-soybean rotation. For this study only six soil samples were collected originating from plots with no fertilizer application and low (0 Mg/ha) and high (16 Mg/ha) mulch rates that had soybeans grown on them.

2.3.4.4 Michigan - KBS Long-Term Ecological Research Station

The Kellogg Biological Station Long-Term Ecological Research project was established in 1987 by Michigan State University and is funded by the National Science Foundation and by the Michigan State University AgBioResearch program. Soil samples were collected from the Main Cropping System Experiment (42.410° N, -85.373° W) which was completed in 1989. The primary soil series is a Kalamazoo loam (fine-loamy, mixed, active, mesic Typic Hapludalfs) with a 2-6 % slope. For the top 15 cm the soil particle size distribution is 32 - 50 % for sand, 34 - 39 % for silt and around 11-19 % for clay (Robertson et al., 2020; Soil Survey Staff, 2019). The soil pH in the 0 to 15 cm ranges from 5.7 to 6.5. The plot size is 87 by 105 m.

It has a factorial randomized complete block design with six replications. The tillage treatments are: (1) conventional chisel (minimum) tillage (MT-Conv); (2) conventional no-tillage (NT-Conv); (3) chisel tillage with reduced- N input (MT-Conv(-

N)); and (4) biologically (organic) based system with chisel tillage (MT-Org) (Martin and Sprunger, 2022; Naasko et al. 2024). The four tillage treatments follow a corn-soybean-wheat rotation, but winter cover crops are incorporated in the reduced input (MT-Org) and biologically based systems (MT-Org) following corn and soybean harvest (corn–ryegrass (*Lolium multiflorum)*–soybean–winter wheat–red clover *(Trifolium pratense)*).

Twenty-four soil samples were collected from the four tillage treatments in 2021.

2.3.5 Surveys and Precipitation Information

The survey used with farmers was designed to study soybean yield gaps due to crop management across the north central US (Edreira et al., 2017). The survey asked questions about crops grown in the past 3 years, tillage, if herbicides or fungicides were used, type of herbicide, if cover cropping, manure rate and type, whether sudden death occurred, drainage system, soybean variety, seed treatment, and weather irrigation were used. Farmers were asked to identify low and high productivity areas in their fields which were sampled separately. Later information was collected on soybean planting and harvest dates. This information was used to determine the field specific precipitation amounts with the help of the Climate Fieldview website (Climate Fieldview, n.d).

The same crop management information was obtained from the LTFS. For the Main Cropping System Experiment (MCSE) in Michigan the management practices and soybean yields were extracted from the publicly available data website and the NSF Long-term Ecological Research Program (DEB 2224712) (Robertson and Snapp, 2019;

Robertson and Simmons, 2020; Martin and Sprunger, 2022).

Recent land management information for the Wooster and Hoytville LTFS specifically in connection to the 2021 soybean yields were provided by Matthew Davis from the OSU agricultural operations department.

Information regarding the 2021 land management history for the East Straw Mulch Experiment, which has been established by the Carbon Management and Sequestration Center (CMASC), was provided by Kyle Sklenka.

2.3.6 Soil Sampling and Processing

With the information from the surveys each individual farm field location was identified. A soil map, and LIDAR elevation information was obtained from the US Soil Survey website and the Ohio Statewide Imagery Program (OSIP). An elevation heatmap was created using a 3D point cloud and mesh processing software CloudCompare. The soil map was overlayed with the elevation heatmap to identify a low and high elevation soil sampling site on each field. Each soil sampling sites was selected based on farmer survey yield information and the premise that the soil units would be identical. The GPS coordinates for both sites were recorded, and the texture specific information was obtained from the US Soil Survey website.

Six to eight soil (0-15 cm depth) cores (2.54 cm dia.) were taken and homogenized to form a composite sample (~1 kg). All cores were taken within a 5 m radius. For the LTFS a randomized soil core sampling was done in a w-shaped pattern. For the Michigan LTFS it was required to sample five predetermined soil sampling subplots. At each subplot two cores (0-15 cm depth; 2.54 cm dia.) were collected and

composited. Soil samples were stored as soon as possible in a cooler with ice and transferred to a -20 °C freezer (Lee et al., 2007; Veum, 2019).

After thawing the soil samples in the 4 °C fridge, the wet soil was sieved to pass a 2 mm mesh size and all organic material, or mineral fragments were removed. A 300 to 500g subsample was air dried for 24 to 48 hours at room temperature, then stored in the 4 °C fridge and used to measure pH, Total C (TC), Total N (TC), soil organic carbon (SOC), and the enzyme activity of β-Glucosidase (GLU), N-Acetyl Glutamate synthase (NAG), and Arylsulfatase (AS). The remaining field moist subsample was stored at -20 ˚C and used for EL-FAME analysis work. Gravimetric water content was determined by weighing before and after a placing a soil subsample in an oven set at 105 ˚C for 24 hours.

2.3.7 Total Nitrogen, Total Carbon, Soil Organic Carbon, and pH

Soil pH was measured with air-dried soils using a 1:1 mixture of soil and deionized water followed by measurement with a glass membrane electrode (Accumet Model 15 pH meter).

Total nitrogen (TN) and total carbon TC was determined on sieved air-dried soil samples that had been crushed with a pestle and mortar to pass a 106 μm sieve (USA Standard Test Sieve Number 104). This subgroup was then used in an elemental analyzer system (Carlo Erba CHN EA 1108, now Thermo Fisher Scientific, Waltham, MA) (Nelson & Sommers, 1996, Matejovic, 1997).

Inorganic carbon (SIC) was determined by placing half of the subsample into a furnace for 16 hours at 450 ˚C (Ball, 1964; Davies, 1974; Ben-Door & Banin, 1989;

Soon and Abboud, 1991; Nelson & Sommers, 1996). Past publications determined that organic matter content by loss-on-ignition at this 400 ˚C temperature resulted in a strong correlation with soil organic carbon content that was determined via wet-oxidation (dichromate) (Ben-Door & Banin, 1989, Nelson & Sommers, 1996). The heating regime of 375 ˚C to 450 ˚C oxidizes all organic matter without creating significant errors due to losses by crystal water or hydroxyl groups from minerals (Davies, 1974; Nelson & Sommers, 1996). After the furnace treatment the subsamples were dry combusted a second time in the elemental analyzer system. SOC was calculated by subtracting the recorded SIC concentration from the TC concentration. In the final step, TN, TC, and the SOC variable, were divided and multiplied by the percentage of clay and separately by the percentage of sand to determine a relationship of these properties with soybean yield.

2.3.8 EL-FAME

The soil microbial community composition was obtained by running the Ester-Linked Fatty Acid Methyl Ester (EL-FAME) analysis which was described by Schutter and Dick (2000) and is based on a method developed by Dr. Rhae Drijber.

Three g of field moist soil was extracted with a 1:1 hexane/methyl-tert butyl ether and Methyl Nonadecanoate mixture that was then vortexed with a 0.2 M methanolic KOH solution. The tube was placed into a water bath at 37 ˚C and incubated for 1h. During this incubation phase the sample was vortexed for 10 seconds every 10 minutes. Afterwards 1.0 M acetic acid is added to establish a pH of 7. In the next step, 10 ml of hexane is added, and the tube is vortex for 60 seconds followed by centrifuging

(1600 rpm for 20 minutes) that partitioned the EL-FAMEs were into the organic phase. The upper, organic phase was removed and evaporated under a stream of N_2 gas. The dried EL-FAME film was dissolved in 1 ml of the internal standard mixture and transferred into a gas chromatograph (GC) for analysis on the 6890N GC (Agilent Technologies).

The GC was equipped with a flame ionization detector that used a fused silica capillary column (25 m × 0.20 mm × 0.33 μm). The system used ultra-high purity H_2 as the carrier gas and the temperature program was ramped from 190 to 285 ˚C at 10 ˚C per minute. The Microbial ID PLFA identification software (MIDI ver.6.2) was used to identify the biomarker and their relative peak areas. The individual biomarkers concentrations (nmol g^{-1} dry soil) were calculated and categorized based on described procedures in the literature (Olsson, et al., 1995; Frostegård & Bååth, 1996; Zelles, 1999; Schutter and Dick, 2002).

Each EL-FAME is described with a nomenclature. The first number clarifies the number of carbon atoms of the fatty acid molecules. It is followed by a colon and a second number which explains the number of double bonds within the molecule. The suffixes "*c*" and "*t*" are used to indicate *Cis* and *trans* isomers. Branched fatty acids are indicated by the prefixes *i* (iso) and *a* (anteiso). Other notations like "*Me*", "*OH*", "*cy*" are used to describe methyl, hydroxy, and cyclopropane groups.

The total FAME concentration (nmol g^{-1} dry soil) was determined by the sum of all identified EL-FAME biomarkers in a soil sample. The sums of individual EL-FAME biomarkers were used to compute broad taxonomic microbial groups such as Gram-positive bacteria (*a*15:0, *i*15:0, *i*16:0, *a*17:0, *i*17:0) (O'Leary and Wilkinson,

1988; Zelles, 1999), Gram-negative bacteria (cy17:0, cy19:0, 16:1ω7c, 17:1ω8c, 18:1ω7c) (Wilkinson, 1988; Tunlid et al., 1989; Kerger, et al., 1986; Haack, et al., 1994, Zelles, 1999), Actinobacteria (10Me16:0, 10Me17:0, 10Me18:0, 10Me19:1ω7c) (Fischer et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Frostegård et al., 1993, Veum et al. 2021), arbuscular mycorrhizal fungi (AMF; 16:1 ω5c) (Nordby et al., 1981; Olsson et al., 1995; Olsson, 1999; Madan et al., 2002), Protozoa (20:3ω6c, 20:4ω6c) (Guckert et al., 1985), and Eukaryotes (21:0, 22:0, 23:0, and 24:0) (Zelles, 1999) (Appendix Table 2).

Additionally soil microbial ratios were calculated, which included the total fungal/bacterial ratio (tFU/BA), fungal/bacterial ration (FU/BA), gram-positive bacteria/gram-negative bacteria ratio (GP/GN), saturated/monounsaturated fatty acid ratio (SAT/MONO), bacterial/total FAME (BA/ToF), cyclopropane fatty acid 17/ 16:1 precursor ratio (Cy17/16; cy17:0/16:1ω7c), and cyclopropane fatty acid 19/ 18:1 precursor ratio (Cy19/18; cy19:0/18:1ω7c). In published studies these ratios were used to interpret microbial community shifts due to stress related conditions (McKinley et al., 2005; Taguchi et al., 1980; Guckert et al., 1986; Kieft et al., 1994, Bossio and Scow, 1998; Moore-Kucera and Dick, 2007).

The tFU/BA ratio was determined with the sum of the saprotrophic fungal and the arbuscular mycorrhizal fungi (AMF) marker (18:1ω9c, 18:2ω6c, and 16:1ω5c) divided by the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c). The FU/BA ratio is calculated in a very similar way with the exception that the AMF biomarker is removed (Frostegård & Bååth, 1996). The GP/GN ratio is calculated with the sum of 5 Gram-positive bacteria divided by 5

Gram-negative bacteria (Frostegård et al., 1993; Zelles et al., 1994). The SAT/MONO ratio was calculated with the sum of 5 saturated fatty acids (14:0, 15:0, 16:0, 17:0, and 18:0) divided by the sum of 7 monounsaturated fatty acids (16:1ω5c, 16:1ω7c, 17:1ω8c, 18:1ω7c, 18:1ω9c, cy17:0, and cy19:0) (McKinley et al., 2005). To determine the BA/ToF ratio the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c) was divided by the total FAME concentration. For the GP/ToF ratio the sum of 5 Gram-positive bacteria was divided by the total FAME concentration. In the final step, all recorded EL-FAME variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand to determine a possible relationship between soybean yields.

2.3.9 Enzyme Activities

The potential enzyme activity of β-glucosidase (B-GLU), Arylsulfatase (AS), and N-Acetyl-β-glutamate synthase (NAG; also known as β-glucosaminidase) were measured for each dry soil sample. These three enzyme activities are involved in the C cycle (B-GLU, NAG), S cycle (AS), and N cycle (NAG) and were determined by conducting well known enzyme assays. The assay procedures have been described elsewhere: B-GLU (Tabatabai, 1994; Dick, 2011), NAG (Parham and Deng, 2000; Dick, 2011), and AS (Tabatabai, 1994; Dick, 2011). For each enzyme assay three replicate samples and one control of 1g of air-dried soil was prepared. Each sample received the corresponding substrate based on the assay protocol before the 1-hour incubation at 37 ˚C started (Table 2.3). For the control the corresponding substrate was added after the reaction was stopped. Enzyme activities are expressed as mg of p-

nitrophenol (PNP) kg^{-1} dry soil h^{-1}.

Additionally, the sum of B-GLU + AS, B-GLU + NAG, AS + NAG, and B-GLU +AS + NAG was determined. Recent studies determined that multi-assay combinations of enzymes are possible and thereby could be used as a new soil health assessment tool across agroecosystems (Acosta-Martínez et al., 2019). Additionally, the ratio of B-GLU / AS, B-GLU / NAG, AS / NAG, (B-GLU +AS) / NAG, (B-GLU + NAG) / AS, and (AS + NAG) / B-GLU were determined. In the final step, all recorded enzyme variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand to determine a possible relationship between soybean yields.

2.3.10 Soybean Yield Sampling

By communicating with farmers during harvest times the beginning maturity (R7) or full maturity growth stage (R8) was determined. The number of soybean plants were counted within a range of 6 to 14 adjacent rows (dependent on row spacing) with a row distance of 5 m. Afterwards around 1.5 kg of soybean plants were randomly collected and counted inside the predefined area. To avoid yield errors due to moisture differences the plants were placed in a 30 °C incubation chamber to dry. In the first-year soybeans seeds were threshed by hand and in the other two years the Agriculex SPT-1A thresher was used. For the Agriculex SPT-1A thresher a total loss of material below 1 % was determined. The soybeans were stored at 4°C in paper bags to avoid the decomposition of proteins and fats. To determine the moisture, protein, and fat content for the individual soybean samples the FOSS Infratec NOVA grain analyzer system was used. By adjusting the measured moisture content to the 13 % moisture content, which is

commonly used for soybeans, and incorporating the recorded stand count, soybean weight, soybean count, and area specific information the soybean yield (kg ha^{-1}) was determined.

2.3.11 Statistical Analysis

Statistical analyses were performed using RStudio, which is an integrated development environment that uses the R programming language and software environment for statistical computing and graphics (R Core Team, 2022; RStudio, 2018). Normality of all data was done with the Shapiro-Wilk test. For variables that were not normal, they were transformed to obtain linearity by one of the following as appropriate for a given variable: x^3, x^2, $\sqrt{x}$, log(x), $1/\sqrt{x}$, 1/x, x^{-2}, or x^{-3}. The R^2 results were then used to sort the variables in a descending order. The coefficient of determination was interpreted based on descriptions by Chan (2003) (Table 2.2). To accomplish this the *tidyverse* and *dplyr* package were used to organize and filter the data set (Wickham et al., 2019). To create graphs the *ggplot2* and *ggpubr* packages were used (Wickham, 2016).

2.4 Results and Discussion

2.4.1 Yield Regression – Data Set

Previous studies found individual variables had very strong to moderate correlations with crop productivity (Nunes et al., 2018; Lorenz et al. 2020; Faé et al., 2020). However, these strong correlations were primarily observed at LTES or when the data set was split into subgroups related by location or years, thereby reducing the samples

size and environmental factors (Nunes et al., 2018; Faé et al., 2020). Lorenz et al. (2020) found a significant but moderate correlation between β-glucosidase and corn yield in a three-year study at on-farm sites across three states. However, this same study found no significant correlation of soybean yield with β-glucosidase or arylsulfatase activities soil organic matter.

The highest possible R-squared (R^2) value of 0.36 was identified for the numerical variable that differentiated between organic and conventional land management practices. It is also the only variable out of the 10 highest that was rated as a moderate relationship. The remaining 9 were identified to be fair (Table 2.3). The first six highest coefficients of determinations were classified into the land management category. Starting with Residue Score ($R^2 = 0.36$), Very Low Residue Coverage ($R^2 = 0.18$), percentage of Residue Coverage ($R^2 = 0.14$), Tillage Score ($R^2 = 0.13$), and Herbicide Score ($R^2 = 0.13$). The numerical classification design for the three scoring variables is provided in Table 2.1. The only three soil properties that made it into the list of the 10 highest coefficients of determination variables were TN * Sand ($R^2 = 0.12$), TC * Sand ($R^2 = 0.12$)., and 19:0 cyclo * Sand ($R^2 = 0.11$). Because none of the ten variables were interpreted to have a strong coefficient of determination ($R^2 = > 0.49$) no individual variable had the potential to predict soybean yields. Consequently, this confirmed what previous studies expressed (Liebig et al., 2001; Lehman et al., 2015; Ghimire et al., 2023).

No single soil property showed a significant relationship with soybean yield likely because of the confounding effects on these measurements due to soil type, seasonal environmental conditions, and variations in crop management among the participating farmers.

The moderate coefficient of determination for the variable that differentiated between organic and conventional land management practices suggested that soybean yield predictions need to be analyzed separately for conventional and organic data sets because they were statistically different from each other (Figure 2.1). There was a significant t-test difference ($P < 0.0001$) between the soybean yields of 4042 kg ha $^{-1}$ (60.11 bsh ac $^{-1}$) for conventional farmers compared to 2290 kg ha $^{-1}$ (34.05 bsh ac $^{-1}$) for organic farmers. The main reason for the lower yields with organic soybeans is that herbicides are not allowed and there mechanical cultivation requires 74 cm (30-inch) row spacing to control weeds.

The regression analysis also found a possible relationship between residue coverage and soybean yields. This observation cannot be explained by the residue coverage alone because crop residues are directly linked to tillage practices. Intensive tillage practices, like moldboard plowing completely turn over the first 20 to 30 cm of soil, and thereby burying crop residues. Compared to no-tillage with no soil disturbance and retention of crop residues. Furthermore, moldboard plowing has been linked to increases in soil erosion and decreased levels of soil organic matter in the topsoil (Doran et al., 1984; Lal, 1997; Veum et al., 2014; Nunes et al., 2018).

A meta-analysis by Nunes et al., (2020) showed that NT increased surface soil organic C, biological activity, soil structure, and the labile C and N fractions of SOM thereby fostering soil microorganism diversity, providing soil organisms with food and a stabilized habitat This means that the visual analysis of crop residue on the surface of a soil can indirectly explain what tillage practices were used on an agricultural field if no mulch was added to the surface. Furthermore, it could be a visual indicator of the health of the soil if the same tillage practices are repeated over several years. Supporting evidence for this is

that the tillage score had 5th highest coefficient of determination value (Table 2.3).

Lopes et al. (2013) found that no tillage in combination with cover crops resulted in higher cumulative crop yields over a period of from 12 - 17 years that correlated with individual soil properties. However, there was poor coefficients of determination of soil properties with individual yearly crop yields.

2.4.2 Yield Regression – Conventional Farming

Because of the statistical difference in soybean yields between conventional and organic farming practices the data set was split into these two management systems and the data reanalyzed. The highest R^2 value of 0.076 was TN with soybean yield. However, this is a poor coefficient of determination. The remaining nine highest variables ranged from 0.073 to 0.063 (Table 2.4). The R^2 values confirmed that no individual land management, environmental, or soil property variable had the potential to predict soybean yields. Additionally, it demonstrated that crop productivity predictions based on individual variables are much more complex at conventional farm sites even when soil health indicators are normalized or expressed per unit of clay or sand. One interesting observation was that 6 out of the 10 highest R^2 in this analysis involved enzyme activity ratios or individual enzyme activities. Because enzyme activity ratios have not been explored in the published literature more research is required. These results could also imply that biochemical processes are more important for soils under conventional land management practices since they use chemical fertilizers.

The second highest R^2 values of 0.073 was determined on soils from fields that had applied manure once in the past three years. A t-test analysis of this variable had a

significant difference (P= 0.006) with an average soybean yield of 4611 kg ha $^{-1}$ (68.57 bsh ac $^{-1}$) for farm fields that applied manure once every 3 years compared to 3932 kg ha $^{-1}$ (58.47 bsh ac $^{-1}$) for farm fields that applied no manure, every year, or twice in a period of 3 years. Lower soybean yields on sites on which manure was applied more frequently could be due to soluble salt concentration in the soil which could result in plant damage, or that the manure was applied during the early soybean growth stage (Shapiro and Kranz, 2005). Other studies have shown that frequent manure application resulted in higher SOC and TN content, and higher soybean yields compared to those that only received chemical fertilizers (Gai et al. 2018; Hoover et al. 2019; Rurangwa et al. 2018; Nguyen et al., 2013).

These observations contradict our results, but they also showed that a beneficial differences in yield would likely become more apparent when manure amendments are used over long periods of time. But because the farmers in this study did not practice the same tillage, cover cropping, planting date routine consistently year to year, it could explain why a frequent manure application did not result in a higher yields. Regional environmental factors like solar radiation and precipitation during the three growing seasons could also explain the lower soybean yields even though manure was applied each year (Faé et al., 2020). Additionally, manure applications could have not followed the "4Rs" for nutrient stewardship (Right Source, Right Rate, Right Time, and Right Place) resulting in runoff events.

2.4.3 Yield Regression – Organic Farming

The highest R^2 value of 0.294 was determined for the 17:0 FAME biomarker

(transformed), followed by the 17:0 10-ME (R^2 = 0.289; transformed), 19:0 cyclo (R^2 = 0.279; transformed), SOC (R^2 = 0.261; transformed), TN (R^2 = 0.248; transformed), TC (R^2 = 0.234; transformed), 22:0 * Clay % (R^2 = 0.234; not transformed), 19:0 cyclo * Clay % (R^2 = 0.233; not transformed), 22:0 / Sand % (R^2 = 0. 232; transformed), and iso 16:0 (R^2 = 0.230; transformed). All 10 variables were interpreted to have fair coefficients of determination (Table 2.5). Once more, even after separating the data set into two subgroups, no individual land management, environmental, or soil property variable resulted in a strong coefficient of determination (R^2 = > 0.49) that could be used to reliably predict soybean yields. One interesting observation was that most of the 10 highest R^2 values were EL-FAME biomarkers, which was very different than the conventional data set. Hypothetically it could imply that organic farming practices rely more on biological processes in the soil compared to conventional farming practices.

The fair coefficients of determination for SOC, and TN in connection to soybean yields for organically managed farms are very likely explained by their relationship to soil organic matter (SOM). SOM is a key component in soils which positively impacts microbial activity, water retention, the accumulation and transfer of nutrients in particular N, P and S, and its ability to decrease the risk of erosion, and sedimentation (Stevenson 1986; Lal, 2003, 2004b; Fageria, 2012). Therefore, SOM becomes a key component for crop production at an organically managed farm especially because synthetic fertilizers are not allowed. Leithold et al. (2015) found that of fresh organic matter inputs are important to maintain SOM under organic farming. Additionally, they suggested that conventional farming benefits from mineralized nitrogen applications through not only increases in crop yields but also in total biomass compared to unfertilized systems.

2.5 Conclusion

The statistical analysis of 521 variables, which include soil properties (biological, chemical, and physical), land management variables, and environmental variables showed that no individual variable was able to predict soybean yield, even when those variables are transformed to fit a normal distribution. The multiplication and normalizing of soil properties by texture, like clay or sand, resulted in no significant improvement for the coefficient of determination (R^2). A deeper analysis of the data set after it was split into organic and conventional land management practices, to account for the statistical difference in soybean yield between the data sets, resulted in the same finding. Overall, this study confirms the existing assumptions that agricultural productivity prediction needs to be based on a combination of variables that account for regional differences in climate, management practices, and soil type. Additionally, the study raises the *unlikely* possibility of a single indicator to predict agricultural productivity in a national scale to be *very unlikely*. Future studies which have the goal of predicting agronomic productivity should focus on multivariate models. Such a model would allow the identification of beneficial sustainable agricultural practices, and management practices that likely increase agricultural productivity.

2.6 References

Acosta-Martínez, V., Pérez-Guzmán, L., & Johnson, J. M. F., 2019. Simultaneous determination of β-glucosidase, β-glucosaminidase, acid phosphomonoesterase, and arylsulfatase activities in a soil sample for a biogeochemical cycling index. Applied Soil Ecology, 142, 72–80. https://doi.org/10.1016/j.apsoil.2019.05.001

Acosta-Martínez, V., Pérez-Guzmán, L., Veum, K. S., Nunes, M. R., & Dick, R. P., 2021. Metabolic Activity– Enzymes. ASSA, CSSA and SSSA, 194–250. https://doi.org/10.1002/9780891189831.ch11

Akoglu, H., 2018. User's guide to correlation coefficients., Turk. J. Emerg. Med., 18, 91–93., https://doi.org/10.1016/j.tjem.2018.08.001

Alves de Castro Lopes, A., Gomes de Sousa, D.M., Chaer, G.M., Bueno dos Reis Junior, F., Goedert, W.J. and de Carvalho Mendes, I., 2013, Interpretation of Microbial Soil Indicators as a Function of Crop Yield and Organic Carbon. Soil Science Society of America Journal, 77: 461-472. https://doi.org/10.2136/sssaj2012.0191

Bagnall, D. K., Rieke, E. L., Morgan, C. L. S., Liptzin, D. L., Cappellazzi, S. B., & Honeycutt, C. W. (2023). A Minimum Suite of Soil Health Indicators for North America. Soil Security, 100084. https://doi.org/10.1016/j.soisec.2023.100084

Ball, D.F., 1964. Loss-on-Ignition as an estimate of organic matter and organic carbon ion non-calcareous soils. Journal of Soil Science, 15: 84-92. https://doi-org.proxy.lib.ohio-state.edu/10.1111/j.1365-2389.1964.tb00247.x

Ben-Dor E. and Banin A., 1989. Determination of organic matter content in arid-zone soils using a simple "loss-on-ignition" method, Communications in Soil Science and Plant Analysis, 20:15-16, 1675 1695, DOI: 10.1080/00103628909368175

Blanco-Canqui, H., Lal, R., 2007. Soil structure and organic carbon relationships following 10 years of wheat straw management in no-till, Soil and Tillage Research, Volume 95, Issues 1–2, Pages 240-254, https://doi.org/10.1016/j.still.2007.01.004

Bossio, D., Scow, K., 1998. Impacts of Carbon and Flooding on Soil Microbial Communities: Phospholipid Fatty Acid Profiles and Substrate Utilization Patterns. Microb Ecol 35, 265–278. https://doi.org/10.1007/s002489900082

Chahal, I., & Van Eerd, L. L. (2019). Quantifying soil quality in a horticultural-cover cropping system. Geoderma, 352, 38–48. https://doi.org/10.1016/j.geoderma.2019.05.039

Chan Y.H., 2003. "Biostatistics 104: correlational analysis." Singap Med J.;44(12):614–619.

Chu, M., Singh, S., Walker, F. R., Eash, N. S., Buschermohle, M. J., Duncan, L. A., & Jagadamma, S. (2019). Soil Health and Soil Fertility Assessment by the Haney Soil Health Test in an Agricultural Soil in West Tennessee. Communications in Soil Science and Plant Analysis, 50(9), 1123–1131. https://doi.org/10.1080/00103624.2019.1604731

Climate Fieldview. (n.d.). [Extracting precipitation amounts for sites located in Ohio and Michigan]. Retrieved between September 2019 until December 2021

Davies, B. E., 1974. Loss-on-Ignition as an Estimate of Soil Organic Matter1. Soil Science Society of America Journal, 38(1), 150. https://doi.org/10.2136/sssaj1974.03615995003800010046x

Deiss, L., Sall, A., Demyan, M. S., & Culman, S. W., 2021. Does crop rotation affect soil organic matter stratification in tillage systems? Soil and Tillage Research, 209, 104932. https://doi.org/10.1016/j.still.2021.104932

Deng, S. P., & Tabatabai, M. A., 1997. Effect of tillage and residue management on enzyme activities in soils: III. Phosphatases and arylsulfatase. Biology and Fertility of Soils, 24(2), 141–146. https://doi.org/10.1007/s003740050222

Dick, R. P., 2011. Methods of Soil Enzymology. In Soil Science Society of America book series. https://doi.org/10.2136/sssabookser9Dick, W. A., 1984. Influence of Long-Term Tillage and Crop Rotation Combinations on Soil Enzyme Activities. Soil Science Society of America Journal, 48(3), 569–574. https://doi.org/10.2136/sssaj1984.03615995004800030020x

Dick, W.A. and D.M. Van Doren, Jr., 1985. Continuous tillage and rotation combinations effects on corn, soybean, and oat yields. Agron. J. 77:459-465.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986a. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Typic Fragiudulf Soil. The Ohio State University, Ohio Agricultural Research and Development Center Wooster, Ohio, p. 34. Research Bulletin ll81.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986b. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Mollic Ochraqualf Soil. Research Bulletin 1180. The Ohio State University Ohio Agricultural Research and Development Center, Wooster, Ohio, 30 p.

Dick, W.A., Triplett, G.B., Van Doren Jr., D.M., 2013. Triplett-Van Doren Long-term Tillage and Crop Rotation Data (Accessed 15 December 2019). Ohio Agricultural Research and Development Center, Wooster, Ohio. https://kb.osu.edu/handle/1811/55716

Doran, J.W., and T. Parkin., 1994. Defining and assessing soil quality. In: J. Doran, D. Coleman, D. Bezdicek, and B. Stewart, editors, Defining soil quality for a sustainable environment. SSSA Special Publications, Minneapolis, MN. p. 3–21. doi:10.2136/sssaspecpub35.c1

Edreira J. I. R., S. Mourtzinis, S.P. Conley, A.C. Roth; I.A. Ciampitti, M. A. Licht, H. Kandel, P.M. Kyveryga, L.E. Lindsey, D.S. Mueller, S.L. Naeve, E. Nafziger, J.E. Specht, J. Stanley; M.J. Staton,

Faé, G. S., Kemanian, A. R., Roth, G. W., White, C., & Watson, J. E., 2020. Soybean yield in relation to environmental and soil properties. European Journal of Agronomy, 118, 126070. https://doi.org/10.1016/j.eja.2020.126070

Fageria, N. K., 2012. Role of Soil Organic Matter in Maintaining Sustainability of Cropping Systems. Communications in Soil Science and Plant Analysis, 43(16), 2063–2113. https://doi.org/10.1080/00103624.2012.697234

Fischer, A., Kroppenstedt, R. M., & Stackebrandt, E., 1983. Molecular-genetic and Chemotaxonomic Studies on Actinomadura and Nocardiopsis. Microbiology, 129(11),

3433–3446. https://doi.org/10.1099/00221287-129-11-3433

Frostegård, Å., Tunlid, A., & Bååth, E., 1993. Phospholipid Fatty Acid Composition, Biomass, and Activity of Microbial Communities from Two Soil Types Experimentally Exposed to Different Heavy Metals. Applied and Environmental Microbiology, 59(11), 3605–3617. https://doi.org/10.1128/aem.59.11.3605-3617.1993

Frostegård, Å., & Bååth, E., 1996. The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. Biology and Fertility of Soils, 22(1-2), 59–65. https://doi.org/10.1007/bf00384433

Gai, X., Liu, H., Liu, J., Zhai, L., Yang, B., Wu, S., Ren, T., Lei, Q., Wang, H., 2018. Long-term benefits of combining chemical fertilizer and manure applications on crop yields and soil carbon and nitrogen stocks in North China Plain. Agricultural Water Management, 208, 384–392. https://doi.org/10.1016/j.agwat.2018.07.002

Ghimire, R., Thapa, V. R., Acosta-Martinez, V., Schipanski, M., Slaughter, L. C., Fonte, S. J., Shukla, M. K., Bista, P., Angadi, S. V., Mikha, M. M., Adebayo, O., & Noble Strohm, T. 2023. Soil Health Assessment and Management Framework for Water-Limited Environments: Examples from the Great Plains of the USA. Soil Systems, 7(1), 22. https://doi.org/10.3390/soilsystems7010022

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D., 1985. Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. Fems Microbiology Letters, 31(3), 147–158. https://doi.org/10.1111/j.1574-6968.1985.tb01143.x

Haack, S.K., Garchow, H., Odelson, D.A., Forney, L.J., Klug, M.J., 1994. Accuracy, reproducibility, and interpretation of fatty acid methyl ester profiles of model bacterial communities. Appl. Environ. Microbiol. 60, 2483-2493.

Haney, R. L., Haney, E. B., Hossner, L. R., & Arnold, J. G., 2006. Development of a New Soil Extractant for Simultaneous Phosphorus, Ammonium, and Nitrate Analysis. Communications in Soil Science and Plant Analysis, 37(11-12), 1511–1523. https://doi.org/10.1080/00103620600709977

Hoover, N. L., Law, J. Y., Long, L. A. M., Kanwar, R. S., Soupir, M. L., 2019. Long-term impact of poultry manure on crop yield, soil and water quality, and crop revenue. Journal of Environmental Management, 252, 109582. https://doi.org/10.1016/j.jenvman.2019.109582

Nawaz, A., Lal, R., Shrestha, R. K., & Farooq, M., 2016. Mulching Affects Soil Properties and Greenhouse Gas Emissions Under Long-Term No-Till and Plough-Till Systems in Alfisol of Central Ohio. Land Degradation & Development, 28(2), 673–681. https://doi.org/10.1002/ldr.2553

Nelson, D.W. and Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter. In Methods of Soil Analysis (eds D.L. Sparks, A.L. Page, P.A. Helmke, R.H. Loeppert, P.N. Soltanpour, M.A. Tabatabai, C.T. Johnston and M.E. Sumner). https://doi.org/10.2136/sssabookser5.3.c34

Kerger, B. D., Nichols, P. D., Antworth, C. P., Sand, W., Bock, E., Cox, J. C., Langworthy, T. A., & White, D. C., 1986. Signature fatty acids in the polar lipids of acid-producing Thiobacillus spp.: Methoxy, cyclopropyl, alpha-hydroxy-cyclopropyl and branched and normal monoenoic fatty acids. FEMS Microbiology Letters, 38(2), 67–77. https://doi.org/10.1016/0378-1097(86)90144-8

Kieft, T.L., Ringelberg D.B., White D.C., 1994.Changesinester-linkedphospholipid fatty acid profiles of subsurface bacteria during starvation and desiccation in porous medium. Appl. Environ. Microbiol.60:3292–3299. DOI: https://doi.org/10.1128/aem.60.9.3292-3299.1994

Kroppenstedt, R., 1985. Fatty acid and menaquinone analysis of actinomycetes and related organisms. In: Goodfellow, M., Minnikin, D.E. (Eds.), Chemical Methods in Bacterial Systematics. SAB Technical Series 20. Academic Press, London, United Kingdom, pp. 173–199.

Lal, R., 1997. Residue management, conservation tillage and soil restoration for mitigating greenhouse effect by CO2-enrichment. Soil and Tillage Research, 43(1-2), 81–107. https://doi.org/10.1016/s0167-1987(97)00036-6

Lal, R., 2003. Global Potential of Soil Carbon Sequestration to Mitigate the Greenhouse Effect. Critical Reviews in Plant Sciences, 22(2), 151–184. https://doi.org/10.1080/713610854

Lal, R., 2004a. Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. Science, 304(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Lal, R., 2004b. Agricultural activities and the global carbon cycle. Nutrient Cycling in Agroecosystems, 70(2), 103–116. https://doi.org/10.1023/b:fres.0000048480.24274.0f

Lehman, R. M., Acosta-Martinez, V., Buyer, J. S., Cambardella, C. A., Collins, H. P., Ducey, T. F., Halvorson, J. J., Jin, V. L., Johnson, J. M. F., Kremer, R. J., Lundgren, J. G., Manter, D. K., Maul, J. E., Smith, J. L., & Stott, D. E. (2015). Soil biology for resilient, healthy soil. Journal of Soil and Water Conservation, 70(1), 12A18A. https://doi.org/10.2489/jswc.70.1.12a

Lee, Y. B., 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), 1299–1305. https://doi.org/10.2136/sssaj2006.0245

Leithold, G., Hülsbergen, K.-J., & Brock, C., 2014. Organic matter returns to soils must be higher under organic compared to conventional farming. Journal of Plant Nutrition and Soil Science, 178(1), 4–12. https://doi.org/10.1002/jpln.201400133

Liebig, M.A., G. Varvel, and J. Doran., 2001. A simple performance-based index for assessing multiple agroecosystem functions. Agron. J. 93(2):313–318. doi:10.2134/agronj2001.932313x

Liptzin, D., Norris, C. E., Cappellazzi, S. B., Bean, G. M., Cope, M., Greub, K. L. H., Rieke, E. L., Tracy, P. W., Aberle, E., Ashworth, A., Bañuelos Tavarez, O., Bary, A. I., Baumhardt, R. L., Borbón Gracia, A., Brainard, D. C., Brennan, J. R., Briones Reyes, D., Bruhjell, D., Carlyle, C. N., & Crawford, J. J. W. (2022). An evaluation of carbon indicators of soil health in long-term agricultural experiments. Soil Biology and Biochemistry, 172, 108708. https://doi.org/10.1016/j.soilbio.2022.108708

Lorenz, N., B. McSpadden Gardener, B., R. Lee, N., Ramsier, C., & P. Dick, R., 2020. Soil Enzyme Activities Associated with Differential Outcomes of Contrasting Approaches to Soil Fertility Management in Corn and Soybean Fields. Applied Ecology and Environmental Sciences, 8(6), 517–525. https://doi.org/10.12691/aees-8-6-26

Madan, R., Pankhurst, C., Hawke, B., & Smith, S., 2002. Use of fatty acids for identification of AM fungi and estimation of the biomass of AM spores in soil. Soil Biology and Biochemistry, 34(1), 125–128. https://doi.org/10.1016/S0038-0717(01)00151-1

Martin, T., Sprunger, C.D., 2022. Sensitive Measures of Soil Health Reveal Carbon Stability Across a Management Intensity and Plant Biodiversity Gradient. Frontiers in Soil Science. Volume 2, doi: 10.3389/fsoil.2022.917885

Matejovic, I., 1997. Determination of carbon and nitrogen in samples of various soils by dry combustion. Communications in Soil Science and Plant Analysis 28, 1499–1511.

McKinley, V. L., Peacock, A. D., & White, D. C., 2005. Microbial community PLFA and PHB responses to ecosystem restoration in tallgrass prairie soils. Soil Biology and Biochemistry, 37(10), 1946–1958. https://doi.org/10.1016/j.soilbio.2005.02.033

Moebius, B.N., H.M. van Es, R.R. Schindelbeck, O.J. Idowu, D.J. Clune, and J.E. Thies., 2007. Evaluation of laboratory-measured soil properties as indicators of soil physical quality. Soil Sci. 172(11):895–912. doi:10.1097/ ss.0b013e318154b520

Moore-Kucera, J., & Dick, R. P., 2007. PLFA Profiling of Microbial Community Structure and Seasonal Shifts in Soils of a Douglas-fir Chronosequence. Microbial Ecology, 55(3), 500–511. https://doi.org/10.1007/s00248-007-9295-1

Naasko, K., Martin, T., Mammana, C., Murray, J., Mann, M., & Sprunger, C., 2023. Soil protein: A key indicator of soil health and nitrogen management. Soil Science Society of America Journal, 88(1), 89–108. https://doi.org/10.1002/saj2.20600

Nguyen, H.Q., Kanwar, R.S., Hoover, N.L., Dixon, P., Hobbs, J., Pederson, C., Soupir, M.L., 2013. Long-term effects of poultry manure application on nitrate leaching in tile drain water. Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 56, 91–101.

Nordby, H. E., Nemec, S., & Nagy, S., 1981. Fatty acids and sterols associated with citrus root mycorrhizae. Journal of Agricultural and Food Chemistry, 29(2), 396–401. https://doi.org/10.1021/jf00104a043

Norris, C. E., Bean, G. M., Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W., 2020. Introducing the North American project to evaluate soil health measurements. Agronomy Journal, 112(4), 3195–3215. https://doi.org/10.1002/agj2.20234

Nunes, M. R., van Es, H. M., Schindelbeck, R., Ristow, A. J., & Ryan, M., 2018. No-till and cropping system diversification improve soil health and crop yield. Geoderma, 328, 30–43. https://doi.org/10.1016/j.geoderma.2018.04.031

Nunes, M. R., Karlen, D. L., Veum, K. S., Moorman, T. B., & Cambardella, C. A., 2020. Biological soil health indicators respond to tillage intensity: A US meta-analysis. Geoderma, 369, 114335. https://doi.org/10.1016/j.geoderma.2020.114335

O'Leary, W. M., and S. G. Wilkinson. 1988. Gram-positive bacteria, p. 117-185. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Olsson, P. A., 1999. Signature fatty acids provide tools for determination of the distribution and interactions of mycorrhizal fungi in soil. 29(4), 303–310. https://doi.org/10.1111/j.1574-6941.1999.tb00621.x

Olsson, P. A., Bååth, E., Jakobsen, I., & Söderström, B., 1995. The use of phospholipid and neutral lipid fatty acids to estimate biomass of arbuscular mycorrhizal fungi in soil. Mycological Research, 99(5), 623–629. https://doi.org/10.1016/s0953-7562(09)80723-5

Parham, J. A., & Deng, S. P., 2000. Detection, quantification and characterization of β-glucosaminidase activity in soil. Soil Biology and Biochemistry, 32(8-9), 1183–1190. https://doi.org/10.1016/s0038-0717(00)00034-1

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio., 2018. RStudio: Integrated development for R. RStudio.

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., Bañuelos Tavarez, O., Bary, A. I., Baumhardt, R. L., Borbón Gracia, A., Brainard, D. C., & Brennan, J. R., 2022. Evaluation of aggregate stability methods for soil health. Geoderma, 428, 116156. https://doi.org/10.1016/j.geoderma.2022.116156

Robertson, G., D. Wall, K. Gross, and K. Renner., 2020. Baseline Spatial Variability Study at the Kellogg Biological Station, Hickory Corners, MI (1988) ver 36. Environmental Data Initiative. https://doi.org/10.6073/pasta/b49a09795d29e0c1149efcc068f3426e

Robertson, G., Snapp S., 2019. Agronomic Yields in Row Crop Agriculture at the Kellogg Biological Station, Hickory Corners, MI (1989 to 2018) ver 39. Environmental Data Initiative. https://doi.org/10.6073/pasta/00c3343378bf55e0727cccecc8a8ff41

Robertson, G., Simmons J., 2020. Main Cropping System Experiment Field Logs and treatment descriptions at the Kellogg Biological Station, Hickory Corners, MI (1988 to 2020) ver 90. Environmental Data Initiative. https://doi.org/10.6073/pasta/e5642426f9200404b6ba3af45ad9c423

Roper, W.R., Osmond, D.L., Heitman, J.L., Wagger, M.G. and Reberg-Horton, S.C., 2017. Soil Health Indicators Do Not Differentiate among Agronomic Management Systems in North Carolina Soils. Soil Science Society of America Journal, 81: 828-843. https://doi.org/10.2136/sssaj2016.12.0400

Rurangwa, E., Vanlauwe, B., & Giller, K. E. (2018). Benefits of inoculation, P fertilizer and manure on yields of common bean and soybean also increase yield of subsequent maize. Agriculture, Ecosystems & Environment, 261, 219–229. https://doi.org/10.1016/j.agee.2017.08.015

Saroa, G. S., & Lal, R., 2003. Soil restorative effects of mulching on aggregation and carbon sequestration in a Miamian soil in central Ohio. Land Degradation & Development, 14(5), 481–493. https://doi.org/10.1002/ldr.569

Schutter, M. E., & Dick, R. P., 2000. Comparison of Fatty Acid Methyl Ester (FAME) Methods for Characterizing Microbial Communities. Soil Science Society of America Journal, 64(5), 1659–1668. https://doi.org/10.2136/sssaj2000.6451659x

Schutter, M., Sandeno, J., and R.P. Dick. 2001. Seasonal, soil type, and alternative management influences on microbial communities of vegetable cropping systems. Biol. Fert. Soils 34: 397-410.

Schutter, M.E., and R.P. Dick. 2002. Microbial community profiles and activities among aggregates of winter fallow and cover-cropped. Soil Sci. Soc. Am. J. 66:142-153

Shapiro, A. C., Kranz, L. W., & Wortmann, S. C., 2005. Salt thresholds for liquid manure applied to corn and soybean. Transactions of the ASAE, 48(3), 1005–1013.https://doi.org/10.13031/2013.18512

Soil Survey Staff., 2019. Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online. Accessed [Jan/2019-2023]

Soon Y. K., Abboud S., 1991. A comparison of some methods for soil organic carbon determination, Communications in Soil Science and Plant Analysis, 22:9-10, 943-954, DOI: 10.1080/00103629109368465

Stevenson F.J. 1986. Cycles of Soil: Carbon, Nitrogen, Phosphorus, Sulfur and Micronutrients. John Wiley and Sons, New York, 380 pp

Tabatabai, M.A. 1994. Soil enzymes. In: R.W. Weaver, J.S. Angle, and P.S. Bottomley, editors, Methods of soil analysis: Part 2. Microbiological and biochemical properties. Vol. 5. Madison, WI: SSSA. p. 775–833.

Taguchi, M., Izui, K., Katsuki, H., 1980. Augmentation of cyclopropane fatty acid synthesis under stringent control in Escherichia coli. J. Biochem. 88, 1879–1882.

Tunlid, A., Hoitink, H.A.J., Low, C. and White, D.C., 1989. Characterization of bacteria that suppress Rhizoctonia damping-off in bark compost media by analysis of fatty acid biomarkers. Applied and environmental microbiology, 55(6), pp.1368-1374.

Veum, K. S., Goyne, K. W., Kremer, R. J., Miles, R. J., & Sudduth, K. A., 2014. Biological indicators of soil quality and soil organic matter characteristics in an agricultural management continuum. Biogeochemistry, 117(1), 81–99. https://doi.org/10.1007/s10533-013-9868-7

Veum, K.S., Lorenz, T. & Kremer, R.J., 2019., Phospholipid Fatty Acid Profiles of Soils under Variable Handling and Storage Conditions. Agronomy Journal, 111: 1090-1096. https://doi.org/10.2134/agronj2018.09.0628

Veum, K. S., Acosta-Martinez, V., R. Michael Lehman, Li, C., Cano, A., & Nunes, M. R., 2021. PLFA and EL-FAME Indicators of Microbial Community Composition. ASSA, CSSA and SSSA, 251–288. https://doi.org/10.1002/9780891189831.ch12

Wickham H., Averick M., Bryan J., Chang W., McGowan L.D., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Pedersen T.L., Miller E., Bache S.M., Müller K., Ooms J., Robinson D., Seidel D.P., Spinu V., Takahashi K., Vaughan D., Wilke C., Woo K., Yutani H., 2019. “Welcome to the tidyverse.”, Journal of Open Source Software, *4*(43), 1686. doi:10.21105/joss.01686

Wickham H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Wilkinson, S. G., 1988. Gram-negative bacteria, p. 299-457. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Zelles, L., Bai, Q. Y., Ma, R. X., Rackwitz, R., Winter, K., & Beese, F., 1994. Microbial biomass, metabolic activity and nutritional status determined from fatty acid patterns and poly-hydroxybutyrate in agriculturally-managed soils. *Soil Biology and Biochemistry*, *26*(4), 439–446. https://doi.org/10.1016/0038-0717(94)90175-9

Zelles, L., 1997. Phospholipid fatty acid profiles in selected members of soil microbial communities. Chemosphere, 35(1-2), 275–294. https://doi.org/10.1016/s0045-6535(97)00155-0

Zelles, L., 1999. Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. Biology and Fertility of Soils, 29(2), 111–129. https://doi.org/10.1007/s003740050533

Table 2.1. Numerical Scoring tables based on residue coverage (spring season), tillage, and herbicide practices.

Variable	***Very High Coverage***	***High Coverage***	***Average Coverage***	***Low Coverage***	***Very low Coverage***
Residue Cover in %	> 60	30 - 60	15 - 30	6 - 15	< 6
Numerical Score	5	4	3	2	1

Variable	***No Tillage***	***Chisel Plow***	***Chisel Plow + Field Cultivator***	***Deep Disk Tillage***	***Moldboard Plow***
Numerical Score	5	4	3	2	1

Variable	***No Herbicide***	***Only Glyphosate***	***Glyphosate + Other Herbicide***	***Dicamba***
Numerical Score	4	3	2	1

Table 2.2. Interpretation of Pearson's and regression coefficients (Adapted from Akoglu, 2018)

Correlation Coefficient (r)		Coefficient of determination (R^2)	Description [a)]
+ 1	- 1	*1*	Perfect
+ 0.9	- 0.9	*0.81*	Very Strong
+ 0.8	- 0.8	*0.64*	Very Strong
+ 0.7	- 0.7	*0.49*	Strong
+ 0.6	- 0.6	0.36	Moderate
+ 0.5	- 0.5	0.25	Fair
+ 0.4	- 0.4	0.16	Fair
+ 0.3	- 0.3	0.09	Fair
+ 0.2	- 0.2	0.04	Poor
+ 0.1	- 0.1	0.01	Poor
0	0	0	None

a) Chen, 2003

Table 2.3. R2 values for the ten untransformed and transformed variables that had the highest relationships with soybean yield using the complete data set (n=153).

Position	Variable	Transformed Data R^2	Interpretation [a)]	Transformed Data		Untransformed Data		Data Type
						R^2	p-value [b)]	
1	Conv. or Organic Farming	0.363	Moderate	*No*	-	0.363	***	Land Management
2	Residue Score	0.175	Fair	*Yes*	$1/x$	0.117	**	Land Management
3	Very Low Coverage (< 6%)	0.138	Fair	*No*	-	0.138	***	Land Management
4	Residue Coverage (%)	0.137	Fair	*Yes*	$\sqrt{x}$	0.061	*	Land Management
5	Tillage Score	0.131	Fair	*Yes*	$1/x$	0.108	**	Land Management
6	Herbicide Score	0.125	Fair	*Yes*	x^3	0.061	*	Land Management
7	Total Nitrogen * Sand %	0.120	Fair	*Yes*	x^{-3}	0.057	*	Soil Property * Sand %
8	Moldboard Plow usage	0.118	Fair	*Yes*	$\sqrt{x}$	0.118	**	Land Management
9	Total Carbon * Sand %	0.115	Fair	*Yes*	x^{-3}	0.052	*	Soil Property * Sand %
10	19:0 cyclo * Sand %	0.107	Fair	*Yes*	$1/x$	0.061	*	FAME-Biomarker * Sand %

a) Interpretation based on Chen, 2003 (Table 2.2)
*b) *** P < 0.000001; ** P < 0.00005; * P < 0.005*

Table 2.4. R^2 values for the ten untransformed and transformed variables that had the highest relationships with soybean yield using the data set from conventionally managed fields (n=123).

Position	Variable	Transformed Data R^2	Interpretation [a)]	Transformed Data		Untransformed Data		Data Type
						R^2	p-value [b)]	
1	Total Nitrogen	0.076	Poor	*Yes*	$1/\sqrt{x}$	0.065	***	Universal Soil Health Indicator
2	Manure usage 1/3 years	0.073	Poor	*Yes*	$\sqrt{x}$	0.073	***	Land Management
3	(NAG / AS)	0.073	Poor	*Yes*	x^{-3}	0.002	NS	Enzyme Ratio
4	((NAG+AS) / B-GLU) * Clay	0.069	Poor	*Yes*	$\log(x)$	0.061	**	Enzyme Ratio * Clay %
5	(NAG / AS) / Sand	0.069	Poor	*Yes*	x^{-3}	0.006	NS	Enzyme Ratio / Sand %
6	Soil Organic Carbon	0.066	Poor	*Yes*	$1/x$	0.054	**	Universal Soil Health Indicator
7	((B-GLU+NAG) / AS) / Sand	0.065	Poor	*Yes*	$1/x$	0.030	NS	Enzyme Ratio / Sand %
8	Crop Rotation Score	0.065	Poor	*No*	-	0.065	***	Land Management
9	(B-GLU+NAG) / AS	0.065	Poor	*Yes*	x^{-2}	0.014	NS	Enzyme Ratio
10	(B-GLU+AS) / Clay	0.063	Poor	*Yes*	$1/\sqrt{x}$	0.032	*	Enzyme Activity / Clay %

a) Interpretation based on Chen, 2003 (Table 2.2)

*b) *** P < 0.005; ** P < 0.01; * P < 0.05; NS - Not Significant*

Table 2.5. R^2 values for the ten untransformed and transformed variables that had the highest relationships with soybean yield using the data set from organically managed fields (n=30).

Position	Variable	Transformed Data R^2	Interpretation [a)]	Transformed Data		Untransformed Data		Data Type
						R^2	p-value [b)]	
1	17:0	0.294	Fair	*Yes*	x^3	0.224	**	FAME-Biomarker
2	17:0 10-ME	0.289	Fair	*Yes*	x^3	0.197	*	FAME-Biomarker
3	19:0 cyclo	0.279	Fair	*Yes*	x^3	0.239	**	FAME-Biomarker
4	Soil Organic Carbon	0.261	Fair	*Yes*	x^2	0.259	**	Universal Soil Health Indicator
5	Total Nitrogen	0.248	Fair	*Yes*	x^2	0.248	**	Universal Soil Health Indicator
6	Total Carbon	0.234	Fair	*Yes*	x^3	0.224	**	Soil Chemical Property
7	22:0 * Clay %	0.233	Fair	*No*	-	0.233	**	FAME-Biomarker * Clay %
8	19:0 cyclo * Clay %	0.233	Fair	*No*	-	0.233	**	FAME-Biomarker * Clay %
9	22:0 / Sand %	0.232	Fair	*Yes*	x^3	0.207	*	FAME-Biomarker / Sand %
10	iso 16:0	0.230	Fair	*Yes*	x^3	0.173	*	FAME-Biomarker

a) Interpretation based on Chen, 2003 (Table 2.2)
*b) ** P < 0.01; * P < 0.05*

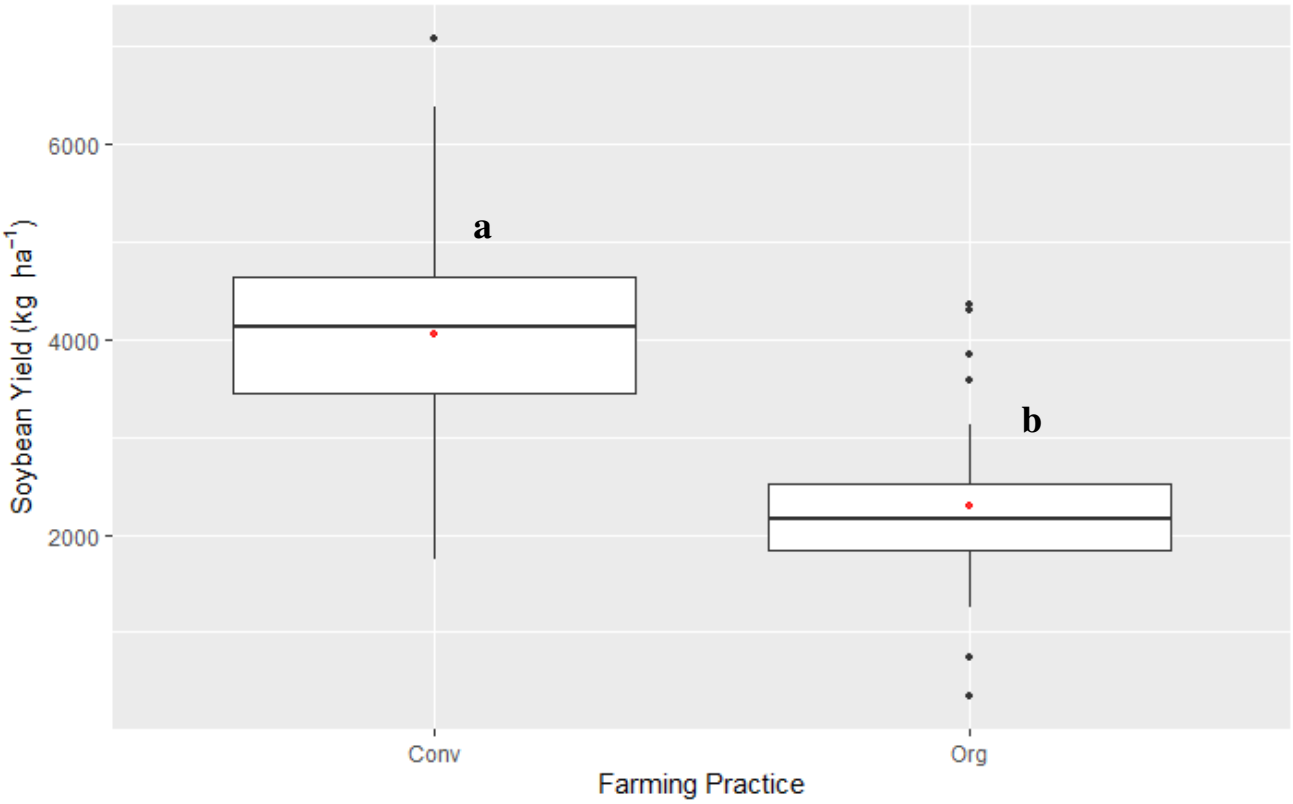


Figure 2.1. Soybean Yields for conventional (n=123) and organic (n=30) farm sites collected from 2019 to 2020. Red point is the mean soybean yield.

Chapter 3: Development of a novel Soil Health Score Index based on a Multivariate Soybean Prediction Model

3.1 Abstract

The statistical analysis in Chapter 2 showed that there were poor correlations of individual soil health measures with soybean yields from farmers' fields in Ohio. The same observation was true for categorical variables related to environmental factors and agricultural management practices. A recent publication by Jemo et al. (2023) showed that it is possible to predict soybean yields with a multivariate model. Therefore, the objective of this study was to develop a robust multivariate soybean yield prediction model that utilized the minimum number of variables selected from the following input variables: enzyme activity, ester-linked fatty acid methyl ester (EL-FAME) biomarkers, total nitrogen (TN), soil organic carbon (SOC), pH, soil texture information, environmental, and land management practice. To develop the data set for model development, soil samples and soybean yields were collected over a period of three years at 153 on-farm sites in Ohio, and for one year at four long-term experimental sites (LTES). The strongest multivariate soybean yield prediction model had an R^2 of 0.86 (99 variables). Cross-validation combined with the complexity reduction algorithm to minimize overfitting and variable intercorrelations, resulted in an optimized model with an R^2 of 0.84 that utilized 77 variables. The regression coefficients for all biological variables were then used to compute the weighted biochemical Soil Health Score (SH) variables.

Furthermore, the sensitivity scores for 521 soil variables at the four LTES were determined using the t-test and Tukey-Kramer post hoc test. This analysis identified two variables (SH-Score $_{[Enz + FAME]}$ and 16:0 iso EL-FAME biomarker) with the highest sensitivity scores for detecting soil/crop management effects.

3.2 Introduction

High quality or healthy soils are essential for optimal agricultural production and delivery of the ecosystem services (Doran, Sarrantonio, & Liebig, 1996). These services are critical for nutrient and water cycles and maintaining and improving biodiversity. But to maintain and improve these functions a quantitative soil health (SH) indicators or models are needed to guide sustainable agricultural practices. This is also important for developing mitigation and adaptation cropping systems to address the negative impacts on crop productivity, water resources and soils of climate change (Lal, 2004a, b; Lal, 2019; Lal 2020).

Because of these risks, the world food charter was adopted that stated, "soil health management is sustainable if the supporting, provisioning, regulating, and cultural services provided by soil are maintained or enhanced without significantly impairing either the soil functions that enable those services or biodiversity" (FAO, 2015). This means that soil degradation should be avoided, and sustainable management should be adopted. But the agricultural sector continues to use chemical fertilizers, tillage, and pesticides, which have been linked to soil organic matter losses (Post and Mann, 1990), and nutrient and pesticide leaching into watersheds (Caraco & Cole, 1999; Mottes et al., 2013). Various sustainable practices can reduce or eliminate these negative environmental impacts such as cover cropping, no-till, and organic amendments.

However, to guide or determine the best cropping system an assessment tool for a SH is needed. The most prominent SH tests are the Haney Soil Health test and the Comprehensive Assessment of Soil Health test (CASH; Cornell Soil Health Laboratory) (Haney et al., 2006; Moebius et al., 2007). However, an investigation by Roper et al. (2017)

showed that they are limited in their ability to differentiate between crop/soil management effects and that they poorly correlate with crop productivity. Other studies conducted in West Tennessee, Pennsylvania, and in Ontario, Canada reported similar deficiencies of these tests (Chu et al., 2019 t; Chahal and Eerd, 2018; Faé et al., 2020,). The tests largely focus on physical and chemical indicators for their SH evaluations and exclude sensitive biological indicators like enzyme assays and microbial properties (Norris et al., 2020).

Conversely arylsulfatase and β-glucosidase have been shown to be sensitive in detecting: (1) cover cropping (Miller and Dick, 1995; Bandick and Dick, 1999; Mendes et al., 1999; Ndiaye et al., 2000; Schutter et al., 2001; Schutter and Dick, 2002; Sprunger et al, 2020), (2) organic amendments (Dungan et al., 2021; Carlson et al., 2015; Yaroshevich, 1966; Khan, 1970; Verstraete and Voets, 1977; Dick et al., 1988; Goyal et al., 1993; Kandeler and Eder, 1993; Werner et al., 1988; Perucci, 1992), (3) heavy metals and herbicides (Hinojosa et al., 2004; Al-Khafaji and Tabatabai; Bardgett et al., 1994; Yeates et al. 1994), (4) climate effects (Acosta-Martínez et a., 2014a,b), (5) tillage (Dick, 1984; Montero et al., 2004; Balota et al., 2004, 2014; Lorenz et al., 2020), and/or (6) perennial management (Vallejo et al., 2009, 2012; Chenhui et a., 2021). For arylsulfatase and β-glucosidase, Bandick and Dick (1999) and Ndiaye et al. (2000) showed that these assays have low in-season variability while demonstrating an overall temporal trajectory, thus enabling the calibration and interpretation of these indicators to detect changes in SH.

To improve the ability to predict yields based on SH indicators, multivariate models are a logical approach, but a review of the literature showed very little research on such model development. Jemo et al. (2023) developed a strong soybean yield prediction model (R^2=0.57) (Table 2.2). The model was based on input variables of soil chemical and

texture properties without any land management variables. Malone et al. (2024) developed a soybean prediction model with a low R^2 of 0.32 using soil data collected from a variety of agricultural fields in Wisconsin using the random forest test. Shukla et al. (2003) used stepwise multiple regression to develop a corn yield prediction model ($R^2 = 0.37$). The Soil Health Institute who collected a large data set of SH measurement from 124 from long-term experimental sites (LTES) across North America which has not been used to develop multivariate yield prediction models (Norris et al. 2020).

In summary very few and less than robust multivariate yield prediction models have been published and individual soil measurements have shown a poor ability to correlate with or predict crop yield (Chapter 2, Roper et al., 2017). Therefore, the objective of this study was to take a different approach in developing a yield prediction model using multi-variate machine learning algorithms (elastic net). Model development was based on data derived from soil analyses and land management variables. Furthermore, various variable transformations and variable ratios were tested to improve model predictability. Lastly the final model with the highest R^2 had to pass cross-validation and penalty tests to reduce overfitting and variable intercorrelations. Soybean was used as the test crop because it is grown worldwide. It was hypothesized that a multi-variate yield prediction model with high R^2 would be developed.

The second objective was to use the optimized model to compute the weighted biochemical Soil Health Score (SHS) variables (derived from EL-FAME and enzyme activities) which would be used for SH assessment or detecting soil/crop management effects - using data from farmer fields and four long-term field experimental sites (LTES). The SHS variable was calculated with a mathematical algorithm that used the biological

regression coefficients from the optimized model. The hypothesis was that by replacing the original biological variables with the SHS variables the adjusted R^2 values would increase.

The third objective was to screen and score 521 soil variables (includes SHS variables) for their ability to detect various soil/crop management treatments at four LTES. The hypothesis was that a minimum subset of variable would have high levels of sensitivity for detecting land management based on means separation analysis.

3.3 Material and Methods

3.3.1 Study Sites

A total of 153 soil samples were collected each spring from 2019 to 2021. Each sample site was identified, and GPS tracked before any sampling occurred. One hundred and six samples came from 18 farms in eight counties in Ohio (Clinton, Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas). In 2021 47 soil samples were collected from four LTES. Three of the LTES are in Ohio and one in Michigan. The soils are classified as: silt loam (59%), as a loam (20%), as a clay loam (11%), and as a silty clay loam (10%).

In most cases there were at least two field where one field had soybeans *(Glycine max)* the first year and the second field had soybeans the following year of sampling. Each farmer was surveyed in person or over the phone to get information on past land management on each field and management plans for each growing season. Five farmers are certified organic farmers and the remaining 13 have conventional land management.

3.3.2 Organic Farm Sites

The organic farm sites (n=18) were in Madison, Handcock, and Clinton county and

have been under agricultural management for 50 to more than 100 years. Fields under organic management had been in place from to one to 20 years and range in size from 5 to 47 ha. Precipitation for each site ranged from 244 to 553 mm (Climate Fieldview, n.d). The growing period from planting to harvest ranged from 120 to 160 days. Most organically managed soils were a Crosby-Lewisburg silt loams (mesic Aeric Epiaqualfs / shallow Aquic Hapludalfs) and four were Mollisols (Soil Survey Staff, 2019). Furthermore, the only four Mollisols in the study were identified at two separate organic farm field locations. All organic farm sites used organic seeds, had a soybean-corn-wheat rotation, a 30-inch (76 cm) row spacing, and no synthetic inputs to meet certified organic standards. However, across the organically, managed fields there was variation in tillage manure applications, and cover cropping.

3.3.3 Conventional Farm Sites

Thirteen conventionally managed fields ranged in size from 4 to 77 ha. Sixteen fields had a soybean-corn-wheat rotation and 72 a soy-corn rotation. Conventional farmers had a 15-inch (38 cm) (n=74) or 7.5-inch (19 cm) (n=12) row spacing. All conventional farmers used synthetic fertilizers and herbicides. Herbicide management was divided in three application categories: only glyphosate (*N-(phosphonomethyl) glycine*), glyphosate with a secondary herbicide, and dicamba (*3,6-dichloro-2-methoxybenzoic acid).* All conventionally managed fields had Alfisols (Soil Survey Staff, 2019). The soil samples came from Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas counties. The agricultural fields under conventional management ranged from 1 to 100 years of usage. One field was converted from native land to farmland. Seasonal

precipitation between planting and harvest was recorded using the Climate Fieldview website which ranged from 329 to 653 mm. (Climate Fieldview, n.d). The individual recorded precipitation levels for conventional farm sites varied between 329 to 653 mm. The growth period ranged from 114 to 173 days. Other practices varied for cover cropping, manure application rates and type, and tillage.

3.3.4 Long Term Field Sites

3.3.4.1 Wooster - Triplett-Van Doren Site

The LTES in Wooster, OH (40.764° N, -81.906° W) was established in 1962 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Wooster silt loam (fine-loamy, mixed, active, mesic Oxyaquic Fragiudalfs) with a 2-6 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25-30 % for sand, 55-60 % for silt and 15% for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). Deiss et al. (2021) reported a range of 5.4 to 6.8 for soil pH.

The experimental has a two-way factorial randomized complete block design with three replications with three tillage treatments, and three crop rotations (Dick and Van Doren Jr., 1985; Deiss et al., 2021). Plot size is 22.3 m by 4.3 m.

The three tillage treatments are: (1) no-tillage (NT); (2) chisel (minimum) tillage (CT); or (3) moldboard plow (MP). The minimum tillage treatment had a para plow from 1962 to 1984, after which a chisel cultivator was used. Chisel tillage loosens the soil and allows up to 30% litter retention on the soil surface. Moldboard tillage inverts soil to a depth of 20 cm and buries the litter, leaving 5 % or less on the soil surface (Dick et al., 2013).

The three crop rotation treatments on the site are: (1) continuous corn *(Zea mays*

L.) (CC); (2) corn and soybean *(Glycine max L.)* (CS); and (3) corn and oat *(Avena sativa L.)* and/or alfalfa *(Medicago sativa)* or clover *(Trifolium repens L.)* (CFF). Nine soil samples were collected in 2021 from the CS rotation plots that grew soybeans.

3.3.4.2 Hoytville - Triplett-Van Doren Site

The LTES in Hoytville, OH (41.222 ° N, -83.762° W) was established in 1963 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Hoytville clay loam (fine, illitic, mesic Mollic Epiaqualfs) with a 0-1 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25 % for sand, 39 % for silt and 36 % for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). In contrast to the Wooster soil, The Hoytville soil has a poor surface and internal drainage, and it cracks when dry. In 1952 a subsurface tile drainage was installed at a depth of 1.2 - 1.4 m (Dick et al., 1986b; Deiss et al., 2021). Deiss et al. (2021) reported a range of 4.3 to 7.5 for soil pH.

It has a two-way factorial randomized complete block design with three replications, and the identical three tillage treatments, and three crop rotations as the Wooster LTES (Dick and Van Doren Jr., 1985; Deiss et al., 2021). The plot size is 30.5 m by 6.4 m. Eight soil samples were collected in 2021 from the CS rotation plots that grew soybeans. (Theoretically 9 but for one plot no soybean yield was recorded).

3.3.4.3 Columbus - Straw Mulch Experiment

The Straw Mulch Experiment (40.017° N, -83.0395° W) was established in 1996 by the Carbon Management and Sequestration Center (CMASC) at the Ohio State University. The objective of this LTES is to determine the effect of wheat straw *(Triticum*

aestivum L.) mulching on soil quality, soil organic carbon (SOC) sequestration and dynamics, and greenhouse gas emissions (Blanco-Canqui and Lal, 2007). No mechanical tillage is used, and glyphosate *(N-Phosphonomethyl glycine)* is used to control weeds. The primary soil series is a Crosby silt loam (fine, mixed, active, mesic Aeric Epiaqualfs) with a 2-6 % slope (Soil Survey Staff, 2019). For the top 15 cm the soil particle size is 22-23 % for sand, 53-56 % for silt, and 22-24 % for clay (Soil Survey Staff, 2019; Nawaz et al., 2016; Saroa and Lal, 2003). Measured soil pH at a depth of 0 to 15 cm ranged from 5.7 to 7.1.

The experimental design is a two-way factorial completely randomized block design (3 replications) with three mulch rates and two fertilizer rates. The fry mulch treatments are: (1) no mulch (control), (2) 8 Mg ha^{-1} yr^{-1}, and (3) 16 Mg ha^{-1} yr^{-1}. The fertilizer treatments are: (1) no fertilizer application (control), or (2) annual broadcast fertilizer application with a rate of 244 kg N ha^{-1} (184 kg N ha^{-1} as Urea) and 60 kg ha^{-1} of NPK). Each year, the wheat straw is applied in the spring followed by fertilizer application in the late spring to early summer. Until 2020 no crops were grown on the plots after which for two years corn and soybean were grown on them. Plot size is 5 by 5 m. Each plot on which the crop experiment took place was separated into two halves (2.5 by 5 m) with a corn-corn and soybean-soybean rotation. For this study only six soil samples were collected originating from plots with no fertilizer application and low (0 Mg/ha) and high (16 Mg/ha) mulch rates that had soybeans grown on them.

3.3.4.4 Michigan - KBS Long-Term Ecological Research Station

The Kellogg Biological Station Long-Term Ecological Research project was established in 1987 by Michigan State University and is funded by the National Science

Foundation and by the Michigan State University AgBioResearch program. Soil samples were collected from the Main Cropping System Experiment (42.410° N, -85.373° W) which was completed in 1989. The primary soil series is a Kalamazoo loam (fine-loamy, mixed, active, mesic Typic Hapludalfs) with a 2-6 % slope. For the top 15 cm the soil particle size distribution is 32 - 50 % for sand, 34 - 39 % for silt and around 11-19 % for clay (Robertson et al., 2020; Soil Survey Staff, 2019). The soil pH in the 0 to 15 cm ranges from 5.7 to 6.5. The plot size is 87 by 105 m.

It has a factorial randomized complete block design with six replications. The tillage treatments are: (1) conventional chisel (minimum) tillage (MT-Conv); (2) conventional no-tillage (NT-Conv); (3) chisel tillage with reduced- N input (MT-Conv(-N)); and (4) biologically (organic) based system with chisel tillage (MT-Org) (Martin and Sprunger, 2022; Naasko et al. 2024). The four tillage treatments follow a corn-soybean-wheat rotation, but winter cover crops are incorporated in the reduced input (MT-Org) and biologically based systems (MT-Org) following corn and soybean harvest (corn–ryegrass (*Lolium multiflorum)*–soybean–winter wheat–red clover *(Trifolium pratense)*).

Twenty-four soil samples were collected from the four tillage treatments in 2021.

3.3.5 Surveys and Precipitation Information

The survey used with farmers was designed to study soybean yield gaps due to crop management across the north central US (Edreira et al., 2017). The survey asked questions about crops grown in the past 3 years, tillage, if herbicides or fungicides were used, type of herbicide, if cover cropping, manure rate and type, whether sudden death occurred, drainage system, soybean variety, seed treatment, and weather irrigation were

used. Farmers were asked to identify low and high productivity areas in their fields which were sampled separately. Later information was collected on soybean planting and harvest dates. This information was used to determine the field specific precipitation amounts with the help of the Climate Fieldview website (Climate Fieldview, n.d).

The same crop management information was obtained from the LTFS. For the Main Cropping System Experiment (MCSE) in Michigan the management practices and soybean yields were extracted from the publicly available data website and the NSF Long-term Ecological Research Program (DEB 2224712) (Robertson and Snapp, 2019; Robertson and Simmons, 2020; Martin and Sprunger, 2022).

Recent land management information for the Wooster and Hoytville LTFS specifically in connection to the 2021 soybean yields were provided by Matthew Davis from the OSU agricultural operations department.

Information regarding the 2021 land management history for the East Straw Mulch Experiment, which has been established by the Carbon Management and Sequestration Center (CMASC), was provided by Kyle Sklenka.

3.3.6 Soil Sampling and Processing

With the information from the surveys each individual farm field location was identified. A soil map, and LIDAR elevation information was obtained from the US Soil Survey website and the Ohio Statewide Imagery Program (OSIP). An elevation heatmap was created using a 3D point cloud and mesh processing software CloudCompare. The soil map was overlayed with the elevation heatmap to identify a low and high elevation soil sampling site on each field. Each soil sampling sites was selected based on farmer survey yield

information and the premise that the soil units would be identical. The GPS coordinates for both sites were recorded, and the texture specific information was obtained from the US Soil Survey website.

Six to eight soil (0-15 cm depth) cores (2.54 cm dia.) were taken and homogenized to form a composite sample (~1 kg). All cores were taken within a 5 m radius. For the LTFS a randomized soil core sampling was done in a w-shaped pattern. For the Michigan LTFS it was required to sample five predetermined soil sampling subplots. At each subplot two cores (0-15 cm depth; 2.54 cm dia.) were collected and composited. Soil samples were stored as soon as possible in a cooler with ice and transferred to a -20 °C freezer (Lee et al., 2007; Veum, 2019).

After thawing the soil samples in the 4 °C fridge, the wet soil was sieved to pass a 2 mm mesh size and all organic material, or mineral fragments were removed. A 300 to 500g subsample was air dried for 24 to 48 hours at room temperature, then stored in the 4 °C fridge and used to measure pH, Total C (TC), Total N (TC), soil organic carbon (SOC), and the enzyme activity of β-Glucosidase (GLU), N-Acetyl Glutamate synthase (NAG), and Arylsulfatase (AS). The remaining field moist subsample was stored at -20 ˚C and used for EL-FAME analysis work. Gravimetric water content was determined by weighing before and after a placing a soil subsample in an oven set at 105 ˚C for 24 hours.

3.3.7 Total Nitrogen, Total Carbon, Soil Organic Carbon, and pH

Soil pH was measured with air-dried soils using a 1:1 mixture of soil and deionized water followed by measurement with a glass membrane electrode (Accumet Model 15 pH meter).

Total nitrogen (TN) and total carbon TC was determined on sieved air-dried soil samples that had been crushed with a pestle and mortar to pass a 106 μm sieve (USA Standard Test Sieve Number 104). This subgroup was then used in an elemental analyzer system (Carlo Erba CHN EA 1108, now Thermo Fisher Scientific, Waltham, MA) (Nelson & Sommers, 1996, Matejovic, 1997).

Inorganic carbon (SIC) was determined by placing the half of the subsample into a furnace for 16 hours at 450 ˚C (Ball, 1964; Davies, 1974; Ben-Door & Banin, 1989; Soon and Abboud, 1991; Nelson & Sommers, 1996). Past publications determined that organic matter content by loss-on-ignition at this 400 ˚C temperature resulted in a strong correlation with soil organic carbon content that was determined via wet-oxidation (dichromate) (Ben-Door & Banin, 1989, Nelson & Sommers, 1996). The heating regime of 375 ˚C to 450 ˚C oxidizes all organic matter without creating significant errors due to losses by crystal water or hydroxyl groups from minerals (Davies, 1974; Nelson & Sommers, 1996). After the furnace treatment the subsamples were dry combusted a second time in the elemental analyzer system. SOC was calculated by subtracting the recorded SIC concentration from the TC concentration. In the final step, TN, TC, and the SOC variable, were divided and multiplied by the percentage of clay and separately by the percentage of sand.

3.3.8 EL-FAME

The soil microbial community composition was obtained by running the Ester-Linked Fatty Acid Methyl Ester method (EL-FAME) which was described by Schutter and Dick (2000) and is based on a method developed by Dr. Rhae Drijber.

Three g of field moist soil was extracted with a 1:1 hexane/methyl-tert butyl ether

and Methyl Nonadecanoate mixture that was then vortexed with a 0.2 M methanolic KOH solution. The tube was placed into a water bath at 37 ˚C and incubated for 1h. During this incubation phase the sample was vortexed for 10 seconds every 10 minutes. Afterwards 1.0 M acetic acid is added to establish a pH of 7. In the next step, 10 ml of hexane is added, and the tube is vortex for 60 seconds followed by centrifuging (1600 rpm for 20 minutes) that partitioned the EL-FAMEs were into the organic phase. The upper, organic phase was removed and evaporated under a stream of N_2 gas. The dried EL-FAME film was dissolved in 1 ml of the internal standard mixture and transferred into a gas chromatograph (GC) for analysis on the 6890N GC (Agilent Technologies).

The GC was equipped with a flame ionization detector that used a fused silica capillary column (25 m × 0.20 mm × 0.33 μm). The system used ultra-high purity H_2 as the carrier gas and the temperature program was ramped from 190 to 285 ˚C at 10 ˚C per minute. The Microbial ID PLFA identification software (MIDI ver.6.2) was used to identify the biomarker and their relative peak areas. The individual biomarkers concentrations (nmol g^{-1} dry soil) were calculated and categorized based on described procedures in the literature (Olsson, et al., 1995; Frostegård & Bååth, 1996; Zelles, 1999; Schutter and Dick, 2002).

Each EL-FAME is described with a nomenclature. The first number clarifies the number of carbon atoms of the fatty acid molecules. It is followed by a colon and a second number which explains the number of double bonds within the molecule. The suffixes "*c*" and "*t*" are used to indicate *Cis* and *trans* isomers. Branched fatty acids are indicated by the prefixes *i* (iso) and *a* (anteiso). Other notations like "*Me*", "*OH*", "*cy*" are used to describe methyl, hydroxy, and cyclopropane groups.

The total FAME concentration (nmol g^{-1} dry soil) was determined by the sum of

all identified EL-FAME biomarkers in a soil sample. The sums of individual EL-FAME biomarkers were used to compute broad taxonomic microbial groups such as Gram-positive bacteria (*a*15:0, *i*15:0, *i*16:0, *a*17:0, *i*17:0) (O'Leary and Wilkinson, 1988; Zelles, 1999), Gram-negative bacteria (cy17:0, cy19:0, 16:1ω7c, 17:1ω8c, 18:1ω7c) (Wilkinson, 1988; Tunlid et al., 1989; Kerger, et al., 1986; Haack, et al., 1994, Zelles, 1999), Actinobacteria (10Me16:0, 10Me17:0, 10Me18:0, 10Me19:1ω7c) (Fischer et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Frostegård et al., 1993, Veum et al. 2021), arbuscular mycorrhizal fungi (AMF; 16:1 ω5c) (Nordby et al., 1981; Olsson et al., 1995; Olsson, 1999; Madan et al., 2002), Protozoa (20:3ω6c, 20:4ω6c) (Guckert et al., 1985), and Eukaryotes (21:0, 22:0, 23:0, and 24:0) (Zelles, 1999) (Appendix Table 2).

Additionally soil microbial ratios were calculated, which included the total fungal/bacterial ratio (tFU/BA), fungal/bacterial ration (FU/BA), gram-positive bacteria/gram-negative bacteria ratio (GP/GN), saturated/monounsaturated fatty acid ratio (SAT/MONO), bacterial/total FAME (BA/ToF), cyclopropane fatty acid 17/ 16:1 precursor ratio (Cy17/16; cy17:0/16:1ω7c), and cyclopropane fatty acid 19/ 18:1 precursor ratio (Cy19/18; cy19:0/18:1ω7c). In published studies these ratios were used to interpret microbial community shifts due to stress related conditions (McKinley et al., 2005; Taguchi et al., 1980; Guckert et al., 1986; Kieft et al., 1994, Bossio and Scow, 1998; Moore-Kucera and Dick, 2007).

The tFU/BA ratio was determined with the sum of the saprotrophic fungal and the arbuscular mycorrhizal fungi (AMF) marker (18:1ω9c, 18:2ω6c, and 16:1ω5c) divided by the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c). The FU/BA ratio is calculated in a very similar way with the

exception that the AMF biomarker is removed (Frostegård & Bååth, 1996). The GP/GN ratio is calculated with the sum of 5 Gram-positive bacteria divided by 5 Gram-negative bacteria (Frostegård et al., 1993; Zelles et al., 1994). The SAT/MONO ratio was calculated with the sum of 5 saturated fatty acids (14:0, 15:0, 16:0, 17:0, and 18:0) divided by the sum of 7 monounsaturated fatty acids (16:1ω5c, 16:1ω7c, 17:1ω8c, 18:1ω7c, 18:1ω9c, cy17:0, and cy19:0) (McKinley et al., 2005). To determine the BA/ToF ratio the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c) was divided by the total FAME concentration. For the GP/ToF ratio the sum of 5 Gram-positive bacteria was divided by the total FAME concentration. In the final step, all recorded EL-FAME variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand.

3.3.9 Enzyme Activity

The potential enzyme activity of β-glucosidase (B-GLU), Arylsulfatase (AS), and N-Acetyl-β-glutamate synthase (NAG; also known as β-glucosaminidase) were measured for each dry soil sample. These three enzyme activities are involved in the C cycle (B-GLU, NAG), S cycle (AS), and N cycle (NAG). The assay procedures have been described elsewhere: B-GLU (Tabatabai, 1994; Dick, 2011), NAG (Parham and Deng, 2000; Dick, 2011), and AS (Tabatabai, 1994; Dick, 2011). For each enzyme assay three replicate samples and one control of 1g of air-dried soil was prepared. Each sample received the corresponding substrate based on the assay protocol before the 1-hour incubation at 37 ˚C started (Table 2.3). For the control the corresponding substrate was added after the reaction was stopped. Color development was measured on a

spectrophotometer. Enzyme activities are expressed as mg of p-nitrophenol (PNP) kg^{-1} dry soil h^{-1}.

Additionally, the sum of B-GLU + AS, B-GLU + NAG, AS + NAG, and B-GLU +AS + NAG was determined. Recent studies determined that multi-assay combinations of enzymes are possible and thereby could be used as a new soil health assessment tool across agroecosystems (Acosta-Martínez et al., 2019). Additionally, the ratio of B-GLU / AS, B-GLU / NAG, AS / NAG, (B-GLU +AS) / NAG, (B-GLU + NAG) / AS, and (AS + NAG) / B-GLU were determined. In the final step, all recorded enzyme variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand to determine a possible relationship between soybean yields.

3.3.10 Soybean Yield Sampling

By communicating with farmers during harvest times the beginning maturity (R7) or full maturity growth stage (R8) was determined. The number of soybean plants were counted within a range of 6 to 14 adjacent rows (dependent on row spacing) with a row distance of 5 m. Afterwards around 1.5 kg of soybean plants were randomly collected and counted inside the predefined area. To avoid yield errors due to moisture differences the plants were placed in a 30 °C incubation chamber to dry. In the first-year soybeans seeds were hand threshed and in the other two years the Agriculex SPT-1A thresher was used. For the Agriculex SPT-1A thresher a total loss of material below 1 % was determined. The soybeans were stored at 4°C in paper bags to avoid the decomposition of proteins and fats. To determine the moisture, protein, and fat content for the individual soybean samples the FOSS Infratec NOVA grain analyzer system was used. By adjusting the measured moisture

content to the 13 % moisture content, which is commonly used for soybeans, and incorporating the recorded stand count, soybean weight, soybean count, and area specific information the soybean yield (kg ha^{-1}) was determined.

3.3.11 Statistical Analysis

Statistical analyses were performed using RStudio, which is an integrated development environment that uses the R programming language and software environment for statistical computing and graphics (R Core Team, 2022; RStudio, 2018). To determine a biochemical Soil Health Score Index with a range of 0 to 1, and simplify the multivariate model interpretation, all measured variables were normalized by dividing the individual measured observation by the maximum recorded observation value in the data set. This rule did not apply to soil properties such as pH (pH_{max} = 12), residue coverage, clay, silt, and sand, which have a scientific or mathematically maximum (100 %) compared to SOC, enzyme activities, or microbial concentrations which can have no fixed maximum. Binary variables (Yes / No) were also transformed into numerical 1 and 0 variables. A list of all variables that were used in this study are listed in Suppl. Table 3.4.

After the normalization steps were completed, the coefficient of determination (R^2) and the adjusted R^2 values for different multivariate regression models was determined by utilizing the *lm* function. By using the *lm* function in R, the statistical analysis is classified as a general linear model. But because the models in this study use continues and binary numeric variables, and that there is no clear answer if a general or generalized linear model should be considered, the *glm* function was also used to analyze the same models in a generalized linear model design. To determine the goodness-of-fit measurement for the

generalized models the McFadden pseudo R^2 values were determined, which cannot be compared with the R^2 results from the general linear model.

To determine the capability of various data clusters to predict soybean yields, separate multivariate models based on soil related properties, and land management characteristics were analyzed. A large variable model design with different combinations of both data clusters were also determined. Because R^2 tends to increase when more variables are added to the model the adjusted R^2 value was used to identify the best multivariate model to account for overfitting (eq. 2). Afterwards the best model was reevaluated through cross-validation and a penalty analysis test. The goal of these tests was to reduce the risk of overfitting and intercorrelations between variables. To achieve this the *cv.glmnet* function was used which was obtained through the *glmnet* package (Friedman et al., 2010; Tay et al., 2023).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_i + \epsilon \qquad \text{(eq. 1)}$$

The resulting multivariate model was then used to extract the coefficient parameters (β_p) for all relevant Enzyme and EL-FAME variables (eq. 1). This information was used to compute the individual coefficient ratios ($\beta_{p\text{-}r}$) separately for all negative and positive coefficient parameters. To compute these ratios all positive and negative coefficients were summed up and each individual coefficient was divided by the computed positive or negative summed value. By multiplying the ratio ($\beta_{p\text{-}r}$) with the corresponding normalized explanatory variable (x_i) and summing the results the positive and negative weighted biochemical Soil Health Scores (SHS) were determined for each observation. To calculate the fitted biochemical SHS (weighted mean), the positive and negative SHS and the coefficient parameters were used (Figure 3.1). With the computed positive, negative, and

fitted SHS a follow up multivariate model analysis was conducted. In this analysis all Enzyme and EL-FAME variables were replaced by computed SHS variables. In this analysis different combinations of variable clusters were explored. The goal was to determine if a simplified model would result in a similar prediction strength (R^2) to the original model and to determine what variable clusters result in drastic changes if they were removed from the model.

Additionally, 521 variables which included the computed SHS variables were statistically analyzed for their capability to detect crop/soil management effects at four LTES. To determine the statistical differentiation power for all variables sensitivity scoring was done. The scoring concept was based on the t-test and Tukey's post-hoc test outcomes (Table 3.1). The p-value for those tests was defined to be below 0.05. To create graphs the *ggplot2* and *cowplot* packages were used (Wickham, 2016).

3.4 Results and Discussion

3.4.1 Multivariate Yield Prediction Models

The multivariate model analysis, based on general and generalized modeling, generated a moderate to strong soybean yield prediction model. For a general model that only included soil property variables and the binary distinction variable between organic and conventional farming practices had an R^2 value of 0.76 ($p<0.0001$) (Table 3.2, Run 24). If the biological variables (enzyme activities and EL-FAME) are excluded from this model the R^2 value drops to 0.46 (Table 3.2, Run 3), which is 30 percentage points lower than the best soil property model. The outcome of this analysis showed that models that include biological properties increase the soybean prediction strength significantly.

The highest R^2 value for a yield prediction model using only land management variables and environmental factors was 0.55 ($p<0.0001$), which is 21 percentage points lower than the model based on soil property variables (Table 3.2, Run 29). But by including texture, soil order, and environmental information, which can be freely extracted from the US Soil Survey and Climate.com website, the model could be increased to an R^2 value of 0.62 ($p<0.0001$) (Table 3.2, Run 31).

The above results showed that in theory a strong or very strong multivariate yield predication models using land management information or soil property data could be developed (Table 2.2). However, the coefficient of determination (R^2) does not indicate causality and cannot account for extreme weather events, which could create outliers that the model cannot predict. The R^2 value also only represents the percentage of variance of the dependent variables and does not correct for the sample size and number of coefficients in a model. Because of these factors the prediction strength of a multivariate model should be based on adjusted R^2 outcomes (eq. 2).

This was done for the two models with the highest R^2 values and resulted in an adjusted R^2 values of 0.59 using soil properties (Table 3.2, Run 24) and 0.53 using land management information and online soil data (Table 3.2, Run 31). These results still support the hypothesis that both model concepts can predict soybean yields, but the results also show that the standard error of the regression is still too large to predict soybean yield accurately.

Separate model runs found that excluding the binary variable which differentiates organic and conventional farming practices resulted in significantly lower R^2 values. It reduced the model by 16 percentage points (Table 3.2, Run 25 vs 24 and 28 vs 31).

The combination of both data sets of variables not only resulted in a higher R^2

value but also in a higher adjusted R^2 value. The multi-variate model that had the highest R^2 value of 0.87 (adjusted R^2 of 0.69, $p<0.0001$) used all 105 explanatory variables (Table 3.2, Run 48). A multi-variate machine learning algorithm (elastic net) was used, which ran a cross-validation (training vs test data) and a complexity model reduction. The outcome of this algorithm was that some variables were eliminated due to variable intercorrelations, resulting in an optimized model with an R^2 value of 0.83.

The same multi-variate machine learning algorithm was used for the second-best model (99 variables) but had the EL-FAME ratio variables excluded (Table 3.2, Run 47). This model had a lower R^2 value of 0.86 ($p<0.0001$) compared to the previous model, but a higher adjusted R^2 of 0.70. After the machine learning algorithm was used the number of variables in the optimized model was reduced to 77 and the R^2 value was 0.84.

Therefore, of the two top models the one with the highest adjusted R^2 value was selected as the optimized model. These results show that the development of crop yield prediction models using the machine learning algorithm method should focus on the adjusted R^2 value rather than the R^2 value to determine the final optimized model.

A study conducted in Nigeria, with soil samples from 350 on-farm sites produced a soybean prediction model with an R^2 of 0.57 (28 variables) (Jemo et al, 2023). This study measured environmental factors, and physical, and chemical soil properties. In this study a random forest machine learning design was used. If biological soil properties had been included the prediction strength of the model could have been higher. This is supported by the model results that only used physical and chemical soil properties and resulted in an R^2 value of 0.46 (Table 3.2, Run 3).

Overall, this statistical model development using various combinations of

variables resulted in a robust soybean yield prediction model adjusted for overfitting applicable Ohio and possibly Michigan. This optimized model was used to determine the weighted biochemical Soil Health Scores (SHS). Additionally, no significant differences between the general and generalized models were found. This observation showed that the interpretation of the multi-variate general models was most reliable and that errors due to non-normal distributions were negligible.

3.4.2 The Biochemical Soil Health Score Index

Individual regression coefficients (β_p) for enzyme activities, enzyme ratios, and EL-FAME variables were extracted from the multi-variate optimized regression soybean prediction model. From this a weighted biochemical Soil Health Score (SHS) with an index range of 0 to 1 was calculated. Separate to this index two different approaches were explored. The first index was based on regression coefficients of individual enzyme activity variables. The second one used regression coefficients based on individual EL-FAME variables (Figure 3.1). To determine these scores the coefficient ratios ($\beta_{p\text{-}f}$) had to be determined and multiplied by the normalized explanatory variables (x_i). An example of the basic computational steps for the determination of the enzymatic SHS is presented in Table 3.3 and Table 3.4. All SHS variables have no unit because each biological variable was divided by the corresponding maximum variable value with the same unit. A linear regression analysis of all nine SHS resulted in a poor correlation with soybean yield. The outcome of this analysis supports the findings in Chapter 2.

3.4.3 Multivariate Model Reevaluation

Statistical model runs where enzyme and EL-FAME variables were replaced by only the fitted SHS variable resulted in the poorest prediction model than those that used the positive and negative SHS variable (Table 3.5, Run 1 vs 2; Run 8 vs 9; Run 15 vs 16). This can be thought of like a balance scale where positive or negative SHS change the position of the scale. But if the fitted SHS is used the balance would stay still.

This is corroborated by the R^2 values for all model runs that used the fitted SHS, and the original large variable run that did not include any enzyme or EL-FAME related variables (Table 3.2, Run 39; Table 3.5, Run 1, 8, 15). All four model runs resulted in a nearly identical R^2 value of 0.65, an adjusted R^2 value of 0.53, and a pseudo R^2 value of 0.65.

Based on this observation any model comparisons were conducted for those models that included the negative and positive SHS. The best model that included the negative and positive SHS, which were developed using the enzymatic assay, enzymatic ratio, and EL-FAME biomarker variables, resulted in a coefficient of determination (R^2) of 0.84 (Table 3.5, Run 16). The adjusted R^2 was determined to be at 0.78, which is significantly larger than the adjusted R^2 (0.70; Table 3.2, Run 47) that was determined for the original model. The increase in the adjusted R^2 was expected since the number of independent predictors dropped from 99 to 48, and the number of observations stayed the same (eq. 2). A correlation graph that plots the predicted soybean yields against the recorded soybean yields is shown in Figure 3.3.

$$Adjusted\ R^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1} \qquad \text{(eq. 2)}$$

Where:

R^2 = Computed R^2 for the model
n = Number of observations
p = Number of independent predictors (variables)

When texture data was excluded from model development the reduction of R^2 values were only ~1 to 3 % (Table 3.5, Run 2 vs 4, Run 9 vs 11, Run 16 vs 18). This observation was unexpected because oil texture is frequently found to be an important factor that controls other soil properties, particularly microbial measurements (Liebig et al., 2001; Vallejo et al., 2009; Dick, 2011; Lehman et al., 2015; Ghimire et al., 2023). This is likely due to the soils used in the model development had similar soils and texture content (Figure 3.4). This needs to be tested on a broad range of soil textures to confirm or refute this observation.

The changes in the coefficient of determination were also small when environmental factors such as precipitation, growing time, and soil order were excluded with small reduction of ~1.28% to 2 % (Table 3.5, Run 2 vs 3, Run 9 vs 10, Run 16 vs 17).

Conversely, if land management variables obtained from farmers or the binary organic and conventional farming variables were removed from the models, R^2 values dropped significantly, ranging from ~25 to 31%; (Table 3.5, Run 2 vs 7, Run 9 vs 14, Run 16 vs 21) and ~21 to 23% (Table 3.5, Run 2 vs 6, Run 9 vs 13, Run 16 vs 20), respectively. This provides evidence for the importance of using crop/soil management information that includes whether organic and conventional farming practices are being followed towards the development of robust yield prediction model.

To reduce cost and labor it is important to have as few measured soil variables as

possible. An analysis found that a soybean yield prediction model based on 3 enzyme assays, survey information, and pH measurements resulted in a strong soybean yield prediction model with an R^2 of 0.71 (Table 3.5, Run 22), which was 6 % higher than the model that used the EL-FAME SHS in the same combination of variables (Table 3.5, Run 23). The difference seems insignificant, however enzyme assays have much less labor and material costs than EL-FAME. First soil handling is easier for enzyme assays, because air-dried soils can be used where as EL-FAME requires soil samples to be stored at -20 °C until analysis. Secondly analytical costs are greater, mainly because EL-FAME requires the use of a gas chromatograph and purchase of MIDI software, whereas enzyme assays only require a spectrophotometer. However, using SHS based on just enzyme variables minimizes the detection of crop/soil management effects as further discussed in section 3.4.5.

Overall, the results showed that it is possible to develop a robust yield prediction model, for soybeans even when texture and environmental information is unavailable. Furthermore, a comprehensive survey with the farmer that captures management practices, has shown to be an important component in improving the model.

3.4.4 EL-FAME, Enzyme activities and ratios

Tillage, at the Hoytville and Wooster long-term field site, did not have any significant effect on the EL-FAME biomarkers for gram-negative bacteria, arbuscular mycorrhizal fungi (AMF), protozoa, fungi/total FAME ratio, AMF/Bacteria ratio, SAT/MONOSAT ratio, the cy17:0/Precursor ratio, β-Glucosidase enzyme activity (β-GLU), N-acetyl-β-glucosaminidase enzyme activity (NAG), β-GLU + NAG, and β-GLU / NAG ratio

(Table 3.6, Table 3.7, Table 3.8).

A significant change in total FAME concentration due to tillage practices was only observed at the Hoytville LTES (Table 3.6). The plots that experienced considerable disturbance through intensive tillage (PT) had approximately an 18 to 20 % lower FAME concentration ($p < 0.05$) compared to MT and NT respectively. This can be explained by the loss of SOC that occurred on the PT plots since the experiment started. In this study SOC concentrations on the PT (14.5 g kg^{-1}) were significantly lower compared to MT (17.05 g kg^{-1}) and NT (17.65 g kg^{-1}) (Table 3.10). Maas et al. (2017) showed similar results in his publication. The moderate correlation between total FAME and SOC ($r = 0.60$) provides support for this (Figure 3.5).

Unexpectedly, total FAME concentration did not differ between tillage treatments at the Wooster LTES. Furthermore, gram-positive bacteria, eukaryotes and the 16:0 iso biomarker for NT were significantly lower than the MT and PT once (Table 3.6). This stands in contrast to many studies where NT is significantly higher than PG for enzyme activities (Acosta-Martínez et al., 2003; Bergstrom et al., 2000; Deng and Tabatabai, 1997; Dick, 1986a; Dick, 1986b), mycorrhizal fungi (Drijber et al., 2000; McGonigle et al., 1999), fungal to bacteria ratios (Frey et al., 1999; Drijber et al., 2000; Helgason et al., 2009), macro aggregation (Kumar et al., 2012a; Six et al., 2000b), TC and TN concentration (Feng et al., 2003; Hendrix et al., 1986), SOC concentration (Kumar et al., 2012a; Dick, 1986a; Dick, 1986b) and available water and field capacity (Kumar et al., 2012a, Kumar et al., 2012b). Other papers also reported changes in the composition of microbial communities in the soils due to the tillage intensity (Jackson et al., 2003; Drijber et al., 2000; Frey et al., 1999; Doran, 1980). At the Wooster site virtually none of the soil enzyme activities or microbial community structure (EL-FAME

biomarkers and total EL-FAME) follow the above findings in the literature.

A potential explanation for the unexpected results for the Wooster LTES could be due to plot renovations in 2006. The objective of this renovation was to remove effects of extensive erosion that occurred after 44 years (W. Dick, 2024, pers. comm.). The erosion created a berm at the bottom of the PT plots where water would accumulate. The accumulation of excess water on PT plots would be affecting yields negatively. In the end the decision was made to relevel the plots and move soil from the plot boarder area. In the following 15 years the soil was very likely redistributed on the PT and CT plots due to reoccurring tillage practices. This could explain the lack of soil management treatment effects.

At the Hoytville site the NT plots resulted in an approximately 23% to 25% lower fungal to bacterial ratios, which raises the question about the it as an indicator for detecting crop/soil management effects (Table 3.7). The AMF showed no significant difference between the tillage practices (Table 3.7) which is likely due to AMF needing a host plant to produce hyphae from the roots which was not there during the spring sampling. Other studies have found elevated AMF level later in the growing season with reduced tillage systems (Feng et al., 2003; Helgason et al., 2009; Mathew et al., 2012, Mbuthia et al., 2015).

The KBS and the Straw Mulch experimental sites have resulted in significant differences between treatments for the AMF concentration (Table 3.6). On the Straw Mulch LTES, the soils for each plot have not been disturbed since the experiment started in 1996 (Jacinthe et al., 2002). This suggests that the amendment of mulch at the soil surface provided the necessary carbohydrates for the microorganisms to thrive. The continuous presence of a protective residue layer against environmental impacts and the establishment of beneficial microclimate likely accelerated this process (Opara-Nadi and Lal, 1987; Duiker and Lal, 2000).

This beneficial environment for the microbial community in the soil very likely explains the significant treatment effects on all EL-FAME biomarkers, and enzyme activities, which suggests that mulch amendments are beneficial for the health of a soil (Table 3.6 Table 3.8).

At the Michigan LTES most EL-FAME biomarkers and enzyme assay were significantly higher for the CSW-MT-Conv$_{(-)}$ and CSW-MT-Org compared to the CSW-MT-Conv and CSW-NT-Conv plots (Table 3.6; Table 3.8). This indicates that continuous cover cropping in a corn-soy-wheat rotation has a beneficial effect on soil microbiology and the biochemical nutrient cycles in soil. This is further demonstrated by the total nitrogen concentrations where treatments are in the order of CSW-MT-CC-Conv$_{(-)}$ > CSW-MT-CC-Org > CSW-MT-Conv = CSW-NT-Conv (Table 3.10).

The individual EL-FAME biomarker 16:0 iso was significantly affected by tillage ($p<0.001$) at the Hoytville LTES following a PT < MT < NT order (Table 3.6). At the Wooster LTES the order was determined to be PT >= MT > NT. This observation is unusual because gram-positive bacteria tend to be much larger in size, have thicker cell walls, and have the capability to resist water stress compared to gram-negative bacteria. It is possible that the 16:0 iso biomarker was wrongly classified as a gram-positive bacteria by O'Leary and Wilkinson (1988) and that it could belong to a gram-negative bacteria or fungal organism. Zelles (1997) suggested that this might be possibility.

For the Michigan and Straw experiment the 16:0 iso provided additional evidence that mulch amendments and cover cropping were beneficial for the health of the soil (Table 3.6). Furthermore, the 16:0 iso biomarker was the most sensitive among all EL-FAME and enzyme activity variables with a sensitivity score of 3.25 (max = 4). A search of the literature found no published information of this biomarker being able to detect treatment effects. This is

likely because EL-FAME biomarkers are frequently organized into taxonomic groups and not as individual biomarkers. This result means that previously conducted studies might have overlooked a sensitive soil health indicator.

When we consider the sensitivity scores of other EL-FAME and enzyme variables some of them (gram-positive *(2.75),* eukaryotes *(2.75),* β-GLU + AS *(2.75),* β-GLU + NAG + AS *(2.75)*) show the potential to detect treatment effects on a local scale and possibly even on a regional one, but not likely on a national scale.

3.4.5 Yield, Soil Health Score, Total Carbon, Soil org. Carbon, and Total Nitrogen

A difference in soybean yield was determined at the Hoytville long-term field site ($p < 0.05$). The plots that have been tilled intensively with a moldboard plow (PT) resulted in a 16 to 21 % lower yields than those that were chisel tilled (MT) or were never tilled (NT) (Table 3.10). Difference between NT and PT soils have been published with NT soils having higher TC and TN concentration (Feng et al., 2003; Hendrix et al., 1986), SOC concentration (Kumar et al., 2012a; Dick, 1986a; Dick, 1986b) and available water and field capacity (Kumar et al., 2012a, Kumar et al., 2012b).

This research indicates that PT results in lower soil quality that in turn reduces yields over time. This corresponds to reduced levels of TC, SOC, TN, and all three SHS variables for PT over the MT and NT (Table 3.10). This increase in soybean yield is inconsistent with what has been reported for the Hoytville site in the past. In past studies the 5-year average yields for the NT plots were always lower compared to the PT plots at the Hoytville long-term field site (Dick and Van Doren, 1985; Dick et al., 1986). But a study conducted from 1988 to 1994 in an adjacent field at the Hoytville site showed that no-till and minimal tillage practices had higher

soybean yields compared to intensive tillage practices (Lal, 1996). It may be that it took longer for soils to improve and increase yields after the earlier yield measurements were made.

The only other difference in soybean yield was at the KBS-LTER site in Michigan, where the biologically based land management system (CSW-MT-C-Org), were significantly lower ($p<0.001$) than the once determined at the other three systems (Table 3.10). In 2021 the soybean yield for the CSW-MT-C-Org treatment were approximately 51% to 53% lower than the other treatments, which is significantly lower than a meta-analysis study of soybean yield reported by Kniss et al. (2016). This meta-analysis study reported that on average soybean yields were 11% to 32% lower for organically over conventional managed farms. These significantly lower yields are because organic soybeans are planted in a 30-inch row spacing. Other factors for lower yields could be due to frequent field cultivation to control weeds, no chemical fertilizers, or the lack of pest control. This reduction in soybean yield for the organic treatment could also be because in 9 of 10 growing seasons (1994-2021), soybeans have been planted last (7 days later on average than other treatments). Singh and Siler (2022) found that row spacing, and late planting date at conventional on-farm fields in Michigan resulted in lower soybean yields. Similar results were found in Ohio which showed that if the seeding rate is not increased with later planting dates lower yields will be lower (Fabiano et al., 2023).

Past soybean yield data for the Michigan LTES was obtained for each treatment. The 1990, 1991, and 1992 years were excluded because soybeans were planted in different growing seasons which would not account for weather related conditions, pest problems and soil nutrient conditions (Table 3.11). Running statistics on this data set showed that over 10 growing seasons soybean yields were significantly different among treatments following an order of CSW-NT-Conv = CSW-NT-CC-$Conv_{(-)}$ => CSW-MT-Conv > CSW-MT-CC-Org

(Table 3.11).

A trendline analysis showed that from 1994 to 2021 the mean soybean yields were increasing for the CSW-MT-Conv, CSW-NT-Conv, CSW-NT-CC-Conv$_{(-)}$, whereas the CSW-MT-CC-Org showed a negative trajectory (Figure 3.2). But a coefficient of determination (R^2) analysis revealed that the variances in these linear trendlines were not strong evidence because the R^2 values ranged from 0.32 to 0.001. These low R^2 are likely due to variations in crop/soil management (e.g. planting date, soil nutrient levels, pests, and weather conditions).

The statistical analysis of TC, SOC, and TN showed that all three are affected by reduced tillage, cover cropping, and soil mulch amendments. The effect of tillage, cover cropping, and mulch amendments on these variables has been shown in previous publications (Feng et al., 2003; Hendrix et al., 1986; Kumar et al., 2012a; Dick, 1986a; Dick, 1986b). The only exception in is the Wooster LTES as discussed above was impacted by the renovation to relevel plots in 2006 that had been severely eroded. A coefficient of determination of 0.85 was determined for SOC and TN (Figure 3.6). Similarly, Liptzin et al. (2022) found a high R^2 value of 0.92 from soils collected from 124 LTES in North America.

The SHS variable based on enzymatic variables was determined to be sensitive to tillage practices and cover cropping practices. At the Hoytville LTES the SH scores were affected by tillage in the order of PT < MT < NT (Table 3.10). At the Michigan LTES treatments with cover crops resulted in significantly higher SH scores. However, it could not differentiate between treatments at the straw mulch LTES. The SHS variable based on EL-FAME variables was overall more sensitive in detecting crop/soil management effects. However, the scores were not as affected by tillage compared to the enzymatic SHS variable. The SHS variable based on enzymatic and EL-FAME variables scored like the iso 16:0

biomarker the highest sensitivity score of 3.25 (max=4). This SHS variable was significantly affected by tillage at the Wooster and Hoytville LTES, by mulch amendments, and by cover cropping at the Michigan LTES. Like the iso 16:0 biomarker it determined an unusual tillage treatment effect at the Wooster LTES with an order of PT <= MT > NT. Again, confirming that the 2006 renovation likely affected the soil.

3.5 Conclusions

From data collected on-farm and long-term experimental multiple multi-variate soybean yield prediction models were successfully developed evidenced by high adjusted $R^2 > 0.5$. These strong prediction models could be developed with the help of soil properties or by using land management survey information. But the large multi=variate model that included both soil property and survey resulted in the highest adjusted R^2 of 0.70 (R^2 = 0.84). This very strong soybean prediction model is significantly stronger than previously published models and R^2 ranging from 0.32 to 0.56 (Malone et al. 2024; Jemo et al., 2023). The model not only allowed for prediction of soybean yields for organic and conventional farmers, but also the creation of an algorithm to calculate a biochemical Soil Health Score (SHS) based on information from three enzyme assays and EL-FAME biomarkers. This SHS was able to quantify crop/soil management effects independent of soil type at four long-term experimental sites much more reliably than the most used SH indicators. The sensitivity scores identified not only the SHS to be the most sensitive SH indicator but also the EL-FAME 16:0 iso biomarker. Further refinement of the model is needed by utilizing data sets generated across diverse soils, agro-ecosystems, and environments and then develop models applicable to other crops such as corn and wheat.

3.6 References

Acosta-Martínez, V., Zobeck, T. M., Gill, T. E., & Kennedy, A. C., 2003. Enzyme activities and microbial community structure in semiarid agricultural soils. Biology and Fertility of Soils, 38(4), 216–227. https://doi.org/10.1007/s00374-003-0626-1

Acosta-Martínez, V., Cotton, J., Gardner, T., Moore-Kucera, J., Zak, J., Wester, D., & Cox, S., 2014a. Predominant bacterial and fungal assemblages in agricultural soils during a record drought/heat wave and linkages to enzyme activities of biogeochemical cycling. Applied Soil Ecology, 84, 69–82. https://doi.org/10.1016/j.apsoil.2014.06.005

Acosta-Martinez, V., Moore-Kucera, J., Cotton, J., Gardner, T., & Wester, D., 2014b. Soil enzyme activities during the 2011 Texas record drought/heat wave and implications to biogeochemical cycling and organic matter dynamics. Applied Soil Ecology, 75, 43–51. https://doi.org/10.1016/j.apsoil.2013.10.008

Acosta-Martínez, V., Pérez-Guzmán, L., & Johnson, J. M. F., 2019. Simultaneous determination of β-glucosidase, β-glucosaminidase, acid phosphomonoesterase, and arylsulfatase activities in a soil sample for a biogeochemical cycling index. Applied Soil Ecology, 142, 72–80. https://doi.org/10.1016/j.apsoil.2019.05.001

Al-Khafaji, A.A. and Tabatabai, M.A., 1979. Effects of trace elements on arylsulfatase activity in soils. Soil Science, 127(3), pp.129-133.

Ball, D.F., 1964. Loss-on-Ignition as an estimate of organic matter and organic carbon ion non-calcareous soils. Journal of Soil Science, 15: 84-92. https://doi-org.proxy.lib.ohio-state.edu/10.1111/j.1365-2389.1964.tb00247.x

Balota, E. L., Kanashiro, M., Colozzi Filho, A., Andrade, D. S., & Dick, R. P., 2004. Soil enzyme activities under long-term tillage and crop rotation systems in subtropical agro-ecosystems. Brazilian Journal of Microbiology, 35(4), 300–306. https://doi.org/10.1590/s1517-83822004000300006

Balota, E. L., Yada, I. F., Amaral, H., Nakatani, A. S., Dick, R. P., & Coyne, M. S. (2013). Long-Term Land use Influences Soil Microbial Biomass P and S, Phoshatase and Arylsulfatase activities, and S Mineralization in a Brazilian Oxisol. Land Degradation & Development, 25(4), 397–406. https://doi.org/10.1002/ldr.2242

Bandick, A. K., & Dick, R. P., 1999. Field management effects on soil enzyme activities. Soil Biology and Biochemistry, 31(11), 1471–1479. https://doi.org/10.1016/s0038-0717(99)00051-6

Bardgett, R.D., Speir, T.W., Ross, D.J., Yeates, G.W. and Kettles, H.A., 1994. Impact of pasture contamination by copper, chromium, and arsenic timber preservative on soil microbial properties and nematodes. *Biology and Fertility of Soils*, *18*, pp.71-79.

Ben-Dor E. and Banin A., 1989. Determination of organic matter content in arid-zone soils using a simple “loss-on-ignition” method, Communications in Soil Science and Plant Analysis, 20:15-16, 1675 1695, DOI: 10.1080/00103628909368175

Bergstrom, D. W., Monreal, C. M., Tomlin, A. D., & Miller, J. J., 2000. Interpretation of soil enzyme activities in a comparison of tillage practices along a topographic and textural gradient. Canadian Journal of Soil Science, 80(1), 71–79. https://doi.org/10.4141/s99-034

Blanco-Canqui, H., Lal, R., 2007. Soil structure and organic carbon relationships following 10 years of wheat straw management in no-till, Soil and Tillage Research, Volume 95, Issues 1–2, Pages 240-254, https://doi.org/10.1016/j.still.2007.01.004

Bossio, D., Scow, K., 1998. Impacts of Carbon and Flooding on Soil Microbial Communities: Phospholipid Fatty Acid Profiles and Substrate Utilization Patterns. Microb Ecol 35, 265–278. https://doi.org/10.1007/s002489900082

Bowman, J. P., Skerratt, J. H., Nichols, P. D., & Sly, L. I., 1991. Phospholipid fatty acid and lipopolysaccharide fatty acid signature lipids in methane-utilizing bacteria. FEMS Microbiology Letters, 85(1), 15–22. https://doi.org/10.1111/j.1574-6968.1991.tb04693.x

Bowman, J. P., Sly, L. I., Nichols, P. D., & Hayward, A. C., 1993. Revised Taxonomy of the Methanotrophs: Description of Methylobacter gen. nov., Emendation of Methylococcus, Validation of Methylosinus and Methylocystis Species, and a Proposal that the Family Methylococcaceae Includes Only the Group I Methanotrophs. International Journal of Systematic Bacteriology, 43(4), 735–753. https://doi.org/10.1099/00207713-43-4-735

Caraco, N.F. and Cole, J.J., 1999. Human impact on nitrate export: an analysis using major world rivers. Ambio, 28(2), pp.167-170.

Carlson, J., Saxena, J., Basta, N. et al., 2015. Application of organic amendments to restore degraded soil: effects on soil microbial properties. Environ Monit Assess 187, 109. https://doi.org/10.1007/s10661-015-4293-0

Chahal, I., & Van Eerd, L. L. (2019). Quantifying soil quality in a horticultural-cover cropping system. Geoderma, 352, 38–48. https://doi.org/10.1016/j.geoderma.2019.05.039

Chu, M., Singh, S., Walker, F. R., Eash, N. S., Buschermohle, M. J., Duncan, L. A., & Jagadamma, S. (2019). Soil Health and Soil Fertility Assessment by the Haney Soil Health Test in an Agricultural Soil in West Tennessee. Communications in Soil Science and Plant Analysis, 50(9), 1123–1131. https://doi.org/10.1080/00103624.2019.1604731

Climate Fieldview. (n.d.). [Extracting precipitation amounts for sites located in Ohio and Michigan]. Retrieved between September 2019 until December 2021

Davies, B. E., 1974. Loss-on-Ignition as an Estimate of Soil Organic Matter1. Soil Science Society of America Journal, 38(1), 150. https://doi.org/10.2136/sssaj1974.03615995003800010046x

Deiss, L., Sall, A., Demyan, M. S., & Culman, S. W., 2021. Does crop rotation affect soil organic matter stratification in tillage systems? Soil and Tillage Research, 209, 104932. https://doi.org/10.1016/j.still.2021.104932

Deng, S. P., & Tabatabai, M. A., 1997. Effect of tillage and residue management on enzyme activities in soils: III. Phosphatases and arylsulfatase. Biology and Fertility of Soils, 24(2), 141–146. https://doi.org/10.1007/s003740050222

Dick, R.P., Rasmussen P.E., & Kerle E.A., 1988. Influence of long-term residue management on soil enzyme activity in relation to soil chemical properties of a wheat-fallow system. Biol. Fert. Soils. 6:159-164.

Dick, R. P., 2011. Methods of Soil Enzymology. In Soil Science Society of America book series. https://doi.org/10.2136/sssabookser9Dick, W. A., 1984. Influence of Long-Term Tillage and Crop Rotation Combinations on Soil Enzyme Activities. Soil Science Society of America Journal, 48(3), 569–574. https://doi.org/10.2136/sssaj1984.03615995004800030020x

Dick, W. A., 1984. Influence of Long-Term Tillage and Crop Rotation Combinations on Soil Enzyme Activities. Soil Science Society of America Journal, 48(3), 569–574. https://doi.org/10.2136/sssaj1984.03615995004800030020x

Dick, W.A. and D.M. Van Doren, Jr., 1985. Continuous tillage and rotation combinations effects on corn, soybean, and oat yields. Agron. J. 77:459-465.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986a. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Typic Fragiudulf Soil. The Ohio State University, Ohio Agricultural Research and Development Center Wooster, Ohio, p. 34. Research Bulletin ll81.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986b. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Mollic Ochraqualf Soil. Research Bulletin 1180. The Ohio State University Ohio Agricultural Research and Development Center, Wooster, Ohio, 30 p.

Dick, W.A., Triplett, G.B., Van Doren Jr., D.M., 2013. Triplett-Van Doren Long-term Tillage and Crop Rotation Data (Accessed 15 December 2019). Ohio Agricultural Research and Development Center, Wooster, Ohio. https://kb.osu.edu/handle/1811/55716

Doran, J. W., Sarrantonio, M., & Liebig, M. A., 1996. Soil Health and Sustainability. Advances in Agronomy, 1–54. https://doi.org/10.1016/s0065-2113(08)60178-9

Doran, J. W., 1980. Microbial Changes Associated with Residue Management with Reduced Tillage. Soil Science Society of America Journal, 44(3), 518–524. https://doi.org/10.2136/sssaj1980.03615995004400030016x

Dowling, N.J.E., Nichols, P.D. and White, D.C., 1988. Phospholipid fatty acid and infra-red spectroscopic analysis of a sulfate-reducing consortium. FEMS Microbiol. Ecol. 53, 325-334.

Dowling, N.J.E., Widdel, F. and White, D.C., 1986. Phospholipid Ester-linked Fatty Acid Biomarkers of Acetate-oxidizing Sulphate-reducers and Other Sulphide-forming Bacteria J. Gen. Microbial. 132, 1815-1829

Drijber, R. A., Doran, J. W., Parkhurst, A. M., & Lyon, D. J., 2000. Changes in soil microbial community structure with tillage under long-term wheat-fallow management. Soil Biology and Biochemistry, 32(10), 1419–1430. https://doi.org/10.1016/s0038-0717(00)00060-2

Duiker, S. W., & Lal, R., 2000. Carbon budget study using CO2 flux measurements from a no till system in central Ohio. Soil and Tillage Research, 54(1-2), 21–30. https://doi.org/10.1016/s0167-1987(99)00101-4

Dungan, R. S., Leytem, A. B., Moore, A. D., Bjorneberg, D. L., Grace, P. R., Brunk, C., & Rowlings, D. W., 2023. Growing and non-growing season nitrous oxide emissions from a manured semiarid cropland soil under irrigation. Agriculture, Ecosystems & Environment, 348, 108413–108413. https://doi.org/10.1016/j.agee.2023.108413

Edreira J. I. R., S. Mourtzinis, S.P. Conley, A.C. Roth; I.A. Ciampitti, M. A. Licht, H. Kandel, P.M. Kyveryga, L.E. Lindsey, D.S. Mueller, S.L. Naeve, E. Nafziger, J.E. Specht, J. Stanley; M.J. Staton,

Fabiano Colet, Lindsey, A. J., & Lindsey, L. E., 2023. Soybean planting date and seeding rate effect on grain yield and profitability. Agronomy Journal, 115(5), 2286–2297. https://doi.org/10.1002/agj2.21434

Faé, G. S., Kemanian, A. R., Roth, G. W., White, C., & Watson, J. E., 2020. Soybean yield in relation to environmental and soil properties. European Journal of Agronomy, 118, 126070. https://doi.org/10.1016/j.eja.2020.126070

FAO. 2015. Status of the World's Soil Resources: Technical Summary. Rome, Italy. FAO. p. 94. http://www.fao.org/3/a-i5126e.pdf

Federle, T. W. 1986. Microbial distribution in soil-new techniques, p. 493-498. In F. Megusar and M. Gantar (ed.), Perspectives in microbial ecology. Slovene Society for Microbiology, Ljubljana, Slovenia.

Feng, Y., Motta, A. C., Reeves, D. W., Burmester, C. H., van Santen, E., & Osborne, J. A., 2003. Soil microbial communities under conventional-till and no-till continuous cotton systems. Soil Biology and Biochemistry, 35(12), 1693–1703. https://doi.org/10.1016/j.soilbio.2003.08.016

Fischer, A., Kroppenstedt, R. M., & Stackebrandt, E., 1983. Molecular-genetic and Chemotaxonomic Studies on Actinomadura and Nocardiopsis. Microbiology, 129(11), 3433–3446. https://doi.org/10.1099/00221287-129-11-3433

Frey, S. D., Elliott, E. T., & Paustian, K., 1999. Bacterial and fungal abundance and biomass in conventional and no-tillage agroecosystems along two climatic gradients. Soil Biology and Biochemistry, 31(4), 573–585. https://doi.org/10.1016/s0038-0717(98)00161-8

Friedman J, Tibshirani R., Hastie T., 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." Journal of Statistical Software, 33(1), 1–22. doi:10.18637/jss.v033.i01.

Frostegård, Å., Tunlid, A., & Bååth, E., 1993. Phospholipid Fatty Acid Composition, Biomass, and Activity of Microbial Communities from Two Soil Types Experimentally Exposed to Different Heavy Metals. Applied and Environmental Microbiology, 59(11), 3605–3617. https://doi.org/10.1128/aem.59.11.3605-3617.1993

Frostegård, Å., & Bååth, E., 1996. The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. Biology and Fertility of Soils, 22(1-2), 59–65. https://doi.org/10.1007/bf00384433

Ghimire, R., Thapa, V. R., Acosta-Martinez, V., Schipanski, M., Slaughter, L. C., Fonte, S. J., Shukla, M. K., Bista, P., Angadi, S. V., Mikha, M. M., Adebayo, O., & Noble Strohm, T. 2023. Soil Health Assessment and Management Framework for Water-Limited Environments: Examples from the Great Plains of the USA. Soil Systems, 7(1), 22. https://doi.org/10.3390/soilsystems7010022

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D., 1985. Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. Fems Microbiology Letters, 31(3), 147–158. https://doi.org/10.1111/j.1574-6968.1985.tb01143.x

Goyal, S., Chander, K., Mundra, M.C. and Kapoor, K.K., 1999. Influence of inorganic fertilizers and organic amendments on soil organic matter and soil microbial properties under tropical conditions. Biology and Fertility of Soils, 29, pp.196-200.

Haack, S.K., Garchow, H., Odelson, D.A., Forney, L.J., Klug, M.J., 1994. Accuracy, reproducibility, and interpretation of fatty acid methyl ester profiles of model bacterial communities. Appl. Environ. Microbiol. 60, 2483-2493.

Haney, R. L., Haney, E. B., Hossner, L. R., & Arnold, J. G., 2006. Development of a New Soil Extractant for Simultaneous Phosphorus, Ammonium, and Nitrate Analysis. Communications in Soil Science and Plant Analysis, 37(11-12), 1511–1523. https://doi.org/10.1080/00103620600709977

Ndiaye, E. L., Sandeno, J. M., McGrath, D., & Dick, R. P., 2000. Integrative biological indicators for detecting change in soil quality. American Journal of Alternative Agriculture, 15(1), 26–36. https://doi.org/10.1017/s0889189300008432

Helgason, B. L., Walley, F. L., & Germida, J. J., 2009. Fungal and Bacterial Abundance in Long-Term No-Till and Intensive-Till Soils of the Northern Great Plains. Soil Science Society of America Journal, 73(1), 120. https://doi.org/10.2136/sssaj2007.0392

Hendrix, P. F., Parmelee, R. W., Crossley, D. A., Coleman, D. C., Odum, E. P., & Groffman, P. M., 1986. Detritus Food Webs in Conventional and No-Tillage Agroecosystems. BioScience, 36(6), 374–380. https://doi.org/10.2307/1310259

Hinojosa, M. B., Carreira, J. A., García-Ruíz, R., & Dick, R. P. (2004). Soil moisture pre-treatment effects on enzyme activities as indicators of heavy metal-contaminated and reclaimed soils. Soil Biology and Biochemistry, 36(10), 1559–1568. https://doi.org/10.1016/j.soilbio.2004.07.003

Jacinthe, P.-A., Lal, R., & Kimble, J. M., 2002. Carbon budget and seasonal carbon dioxide emission from a central Ohio Luvisol as influenced by wheat residue amendment. Soil and Tillage Research, 67(2), 147–157. https://doi.org/10.1016/s0167-1987(02)00058-2

Jackson, L. E., Calderon, F. J., Steenwerth, K. L., Scow, K. M., & Rolston, D. E., 2003. Responses of soil microbial processes and community structure to tillage events and implications for soil quality. Geoderma, 114(3-4), 305–317. https://doi.org/10.1016/s0016-7061(03)00046-6

Jemo, M., K. P. Devkota, T. E. Epule, T. Chfadi, R. Moutiq, Hafidi, M., Silatsa, F. B. T., & J. M. Jibrin., 2023. Exploring the potential of mapped soil properties, rhizobium inoculation, and phosphorus supplementation for predicting soybean yield in the savanna areas of Nigeria. Frontiers in Plant Science, 14. https://doi.org/10.3389/fpls.2023.1120826

Kandeler, E. and Eder, G., 1993. Effect of cattle slurry in grassland on microbial biomass and on activities of various enzymes. Biology and Fertility of Soils, 16, pp.249-254.

Kerger, B. D., Nichols, P. D., Antworth, C. P., Sand, W., Bock, E., Cox, J. C., Langworthy, T. A., & White, D. C., 1986. Signature fatty acids in the polar lipids of acid-producing Thiobacillus spp.: Methoxy, cyclopropyl, alpha-hydroxy-cyclopropyl and branched and normal monoenoic fatty acids. FEMS Microbiology Letters, 38(2), 67–77. https://doi.org/10.1016/0378-1097(86)90144-8

Khan, S., 1970. Enzymatic activity in a gray wooded soil as influenced by cropping systems and fertilizers. Soil Biology and Biochemistry/Soil Biology & Biochemistry, 2(2), 137–139. https://doi.org/10.1016/0038-0717(70)90016-7

Kieft,T.L., Ringelberg D.B., White D.C..1994.Changesinester-linkedphospholipid fatty acid profiles of subsurface bacteria during starvation anddesiccation in porous medium. Appl. Environ. Microbiol.60:3292–3299. DOI: https://doi.org/10.1128/aem.60.9.3292-3299.1994

Kieft, T. L., Wilch, E., O'connor, K., Ringelberg, D. B., & White, D. C., 1997. Survival and phospholipid Fatty Acid profiles of surface and subsurface bacteria in natural sediment microcosms. Applied and Environmental Microbiology, 63(4), 1531–1542. https://doi.org/10.1128/aem.63.4.1531-1542.1997

Kniss, A. R., Savage, S. D., & Jabbour, R., 2016. Commercial Crop Yields Reveal Strengths and Weaknesses for Organic Agriculture in the United States. PLOS ONE, 11(8), e0161673. https://doi.org/10.1371/journal.pone.0161673

Kroppenstedt, R., 1985. Fatty acid and menaquinone analysis of actinomycetes and related organisms. In: Goodfellow, M., Minnikin, D.E. (Eds.), Chemical Methods in Bacterial Systematics. SAB Technical Series 20. Academic Press, London, United Kingdom, pp. 173–199.

Kumar, S., Kadono, A., Lal, R., & Dick, W., 2012a. Long-Term No-Till Impacts on Organic Carbon and Properties of Two Contrasting Soils and Corn Yields in Ohio. Soil Science Society of America Journal, 76(5), 1798–1809. https://doi.org/10.2136/sssaj2012.0055

Kumar, S., Kadono, A., Lal, R., & Dick, W., 2012b. Long-Term Tillage and Crop Rotations for 47–49 Years Influences Hydrological Properties of Two Soils in Ohio. Soil Science Society of America Journal, 76(6), 2195–2207. https://doi.org/10.2136/sssaj2012.0098

Lal, R., 1996. Axle load and tillage effects on crop yields on a Mollic Ochraqualf in Northwest Ohio. Soil and Tillage Research, 37(2-3), 143–160. https://doi.org/10.1016/0167-1987(95)01004-1

Lal, R., 2004a. Agricultural activities and the global carbon cycle. Nutrient Cycling in Agroecosystems, 70(2), 103–116. https://doi.org/10.1023/b:fres.0000048480.24274.0f

Lal, R., 2004b. Soil carbon sequestration to mitigate climate change. Geoderma, 123(1-2), 1–22. https://doi.org/10.1016/j.geoderma.2004.01.032

Lal, R., 2019. Rights-of-Soil. Journal of Soil and Water Conservation, 74(4), 81A86A. https://doi.org/10.2489/jswc.74.4.81a

Lal, R., 2020. Regenerative agriculture for food and climate. Journal of Soil and Water Conservation, 75(5), jswc.2020.0620A. https://doi.org/10.2489/jswc.2020.0620a

Lehman, R. M., Acosta-Martinez, V., Buyer, J. S., Cambardella, C. A., Collins, H. P., Ducey, T. F., Halvorson, J. J., Jin, V. L., Johnson, J. M. F., Kremer, R. J., Lundgren, J. G., Manter, D. K., Maul, J. E., Smith, J. L., & Stott, D. E. (2015). Soil biology for resilient, healthy soil. Journal of Soil and Water Conservation, 70(1), 12A18A. https://doi.org/10.2489/jswc.70.1.12a

Lee, Y. B., 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), 1299–1305. https://doi.org/10.2136/sssaj2006.0245

Liebig, M.A., G. Varvel, and J. Doran., 2001. A simple performance-based index for assessing multiple agroecosystem functions. Agron. J. 93(2):313–318. doi:10.2134/agronj2001.932313x

Liptzin, D., Norris, C. E., Cappellazzi, S. B., Bean, G. M., Cope, M., Greub, K. L. H., Rieke, E. L., Tracy, P. W., Aberle, E., Ashworth, A., Bañuelos Tavarez, O., Bary, A. I., Baumhardt, R. L., Borbón Gracia, A., Brainard, D. C., Brennan, J. R., Briones Reyes, D., Bruhjell, D., Carlyle, C. N., & Crawford, J. J. W. (2022). An evaluation of carbon indicators of soil health in long-term agricultural experiments. Soil Biology and Biochemistry, 172, 108708. https://doi.org/10.1016/j.soilbio.2022.108708

Liptzin, D., Rieke, E. L., Cappellazzi, S. B., G. Mac Bean, Cope, M., Kelsey, et al., 2023. An evaluation of nitrogen indicators for soil health in long-term agricultural experiments. Soil Science Society of America Journal, 87(4), 868–884. https://doi.org/10.1002/saj2.20558

Lorenz, N., B. McSpadden Gardener, B., R. Lee, N., Ramsier, C., & P. Dick, R., 2020. Soil Enzyme Activities Associated with Differential Outcomes of Contrasting Approaches to Soil Fertility Management in Corn and Soybean Fields. Applied Ecology and Environmental Sciences, 8(6), 517–525. https://doi.org/10.12691/aees-8-6-26

Maas, E. D. v.L., Lal, R., Coleman, K., Montenegro, A., & Dick, W. A., 2017. Modeling soil organic carbon in corn (Zea mays L.)-based systems in Ohio under climate change. Journal of Soil and Water Conservation, 72(3), 191–204. https://doi.org/10.2489/jswc.72.3.191

Madan, R., Pankhurst, C., Hawke, B., & Smith, S., 2002. Use of fatty acids for identification of AM fungi and estimation of the biomass of AM spores in soil. Soil Biology and Biochemistry, 34(1), 125–128. https://doi.org/10.1016/S0038-0717(01)00151-1

Malone, L. C., Ruark, M. D., Kucharik, C. J., Whitman, T., & Conley, S. P., 2023. Linking soil health indicators to management history and soybean yield. Field Crops Research, 297, 108951. https://doi.org/10.1016/j.fcr.2023.108951

Martin, T., Sprunger, C.D., 2022. Sensitive Measures of Soil Health Reveal Carbon Stability Across a Management Intensity and Plant Biodiversity Gradient. Frontiers in Soil Science. Volume 2, doi: 10.3389/fsoil.2022.917885

Matejovic, I., 1997. Determination of carbon and nitrogen in samples of various soils by dry combustion. Communications in Soil Science and Plant Analysis 28, 1499–1511.

Mathew, R. P., Feng, Y., Githinji, L., Ankumah, R., & Balkcom, K. S., 2012. Impact of No-Tillage and Conventional Tillage Systems on Soil Microbial Communities. Applied and Environmental Soil Science, 2012, 1–10. https://doi.org/10.1155/2012/548620

Mbuthia, L. W., Acosta-Martínez, V., DeBruyn, J., Schaeffer, S., Tyler, D., Odoi, E., et al., 2015. Long term tillage, cover crop, and fertilization effects on microbial community structure, activity: Implications for soil quality. Soil Biology and Biochemistry, 89, 24–34. https://doi.org/10.1016/j.soilbio.2015.06.016

McGonigle, T. P., Miller, M. H., & Young, D., 1999. *Plant and Soil*, *210*(1), 33–42. https://doi.org/10.1023/a:1004633512450

McKinley, V. L., Peacock, A. D., & White, D. C., 2005. Microbial community PLFA and PHB responses to ecosystem restoration in tallgrass prairie soils. Soil Biology and Biochemistry, 37(10), 1946–1958. https://doi.org/10.1016/j.soilbio.2005.02.033

Mendes, I. C., Anna Katrina Bandick, Dick, R. P., & Bottomley, P. J., 1999. Microbial Biomass and Activities in Soil Aggregates Affected by Winter Cover Crops. Soil Science Society of America Journal, 63(4), 873–881. https://doi.org/10.2136/sssaj1999.634873x

Miller, M., & Dick, R. P., 1995. Dynamics of soil C and microbial biomass in whole soil and aggregates in two cropping systems. Applied Soil Ecology, 2(4), 253–261. https://doi.org/10.1016/0929-1393(95)00060-6

Moebius-Clune, B.N., H.M. van Es, R.R. Schindelbeck, O.J. Idowu, D.J. Clune, and J.E. Thies., 2007. Evaluation of laboratory-measured soil properties as indicators of soil physical quality. Soil Sci. 172(11):895–912. doi:10.1097/ ss.0b013e318154b520

Montero, F. A., Sagardoy, M. A., & Dick, R. P., 2004. Temporal Variability of Microbial Populations and Enzyme Activities of No-tillage Soils in Argentina. Arid Land Research and Management/Arid Land Research and Management, 18(3), 201–215. https://doi.org/10.1080/15324980490451294

Moore-Kucera, J., & Dick, R. P., 2007. PLFA Profiling of Microbial Community Structure and Seasonal Shifts in Soils of a Douglas-fir Chronosequence. Microbial Ecology, 55(3), 500–511. https://doi.org/10.1007/s00248-007-9295-1

Mottes, C., Lesueur-Jannoyer, M., Le Bail, M., & Malézieux, E., 2013. Pesticide transfer models in crop and watershed systems: a review. Agronomy for Sustainable Development, 34(1), 229–250. https://doi.org/10.1007/s13593-013-0176-3

Nawaz, A., Lal, R., Shrestha, R. K., & Farooq, M., 2016. Mulching Affects Soil Properties and Greenhouse Gas Emissions Under Long-Term No-Till and Plough-Till Systems in Alfisol of Central Ohio. Land Degradation & Development, 28(2), 673–681. https://doi.org/10.1002/ldr.2553

Nelson, D.W. and Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter. In Methods of Soil Analysis (eds D.L. Sparks, A.L. Page, P.A. Helmke, R.H. Loeppert, P.N. Soltanpour, M.A. Tabatabai, C.T. Johnston and M.E. Sumner). https://doi.org/10.2136/sssabookser5.3.c34

Nichols, P. D., Smith Glen A, Antworth, C. P., Hanson, R. S., & White, D. C., 1985. Phospholipid and lipopolysaccharide normal and hydroxy fatty acids as potential signatures for methane-oxidizing bacteria. FEMS Microbiology Letters, 31(6), 327–335. https://doi.org/10.1111/j.1574-6968.1985.tb01168.x

Nordby, H. E., Nemec, S., & Nagy, S., 1981. Fatty acids and sterols associated with citrus root mycorrhizae. Journal of Agricultural and Food Chemistry, 29(2), 396–401. https://doi.org/10.1021/jf00104a043

Norris, C. E., Bean, G. M., Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W., 2020. Introducing the North American project to evaluate soil health measurements. Agronomy Journal, 112(4), 3195–3215. https://doi.org/10.1002/agj2.20234

O'Leary, W. M., and S. G. Wilkinson. 1988. Gram-positive bacteria, p. 117-185. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Olsson, P. A., Bååth, E., Jakobsen, I., & Söderström, B., 1995. The use of phospholipid and neutral lipid fatty acids to estimate biomass of arbuscular mycorrhizal fungi in soil. Mycological Research, 99(5), 623–629. https://doi.org/10.1016/s0953-7562(09)80723-5

Olsson, P. A., 1999. Signature fatty acids provide tools for determination of the distribution and interactions of mycorrhizal fungi in soil. 29(4), 303–310. https://doi.org/10.1111/j.1574-6941.1999.tb00621.x

Opara-Nadi, O. A., & Lal, R., 1987. Effects of no-till and disc plowing with and without residue mulch on tropical root crops in southeastern Nigeria. Soil and Tillage Research, 9(3), 231–240. https://doi.org/10.1016/0167-1987(87)90072-9

Parham, J. A., & Deng, S. P., 2000. Detection, quantification and characterization of β-glucosaminidase activity in soil. Soil Biology and Biochemistry, 32(8-9), 1183–1190. https://doi.org/10.1016/s0038-0717(00)00034-1

Parker, J. H., Smith, G. A., Fredrickson, H. L., J. Robie Vestal, & White, D. C., 1982. Sensitive assay, based on hydroxy fatty acids from lipopolysaccharide lipid A, for Gram-negative bacteria in sediments. Applied and Environmental Microbiology, 44(5), 1170–1177. https://doi.org/10.1128/aem.44.5.1170-1177.1982

Parkes, R. J., Dowling, N. J. E., White, D. C., Herbert, R. A., & Gibson, G., 1993. Characterization of sulphate-reducing bacterial populations within marine and estuarine sediments with different rates of sulphate reduction. FEMS Microbiology Ecology, 102, 235–250. https://centaur.reading.ac.uk/34143/

Perucci, P., 1992. Enzyme activity and microbial biomass in a field soil amended with municipal refuse. Biology and Fertility of soils, 14, pp.54-60.

Post, W. M., & Mann, L. K., 1990. Changes in Soil Organic Carbon and Nitrogen as a Result of Cultivation. Www.osti.gov. https://www.osti.gov/biblio/1389523

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio., 2018. RStudio: Integrated development for R. RStudio.

Robertson, G., D. Wall, K. Gross, and K. Renner., 2020. Baseline Spatial Variability Study at the Kellogg Biological Station, Hickory Corners, MI (1988) ver 36. Environmental Data Initiative. https://doi.org/10.6073/pasta/b49a09795d29e0c1149efcc068f3426e

Robertson, G., Snapp S., 2019. Agronomic Yields in Row Crop Agriculture at the Kellogg Biological Station, Hickory Corners, MI (1989 to 2018) ver 39. Environmental Data Initiative. https://doi.org/10.6073/pasta/00c3343378bf55e0727cccecc8a8ff41

Robertson, G., Simmons J., 2020. Main Cropping System Experiment Field Logs and treatment descriptions at the Kellogg Biological Station, Hickory Corners, MI (1988 to 2020) ver 90. Environmental Data Initiative. https://doi.org/10.6073/pasta/e5642426f9200404b6ba3af45ad9c423

Roper, W.R., Osmond, D.L., Heitman, J.L., Wagger, M.G. and Reberg-Horton, S.C., 2017. Soil Health Indicators Do Not Differentiate among Agronomic Management Systems in North Carolina Soils. Soil Science Society of America Journal, 81: 828-843. https://doi.org/10.2136/sssaj2016.12.0400

Saroa, G. S., & Lal, R., 2003. Soil restorative effects of mulching on aggregation and carbon sequestration in a Miamian soil in central Ohio. Land Degradation & Development, 14(5), 481–493. https://doi.org/10.1002/ldr.569

Schutter, M. E., & Dick, R. P., 2000. Comparison of Fatty Acid Methyl Ester (FAME) Methods for Characterizing Microbial Communities. Soil Science Society of America Journal, 64(5), 1659–1668. https://doi.org/10.2136/sssaj2000.6451659x

Schutter, M.E., Sandeno, J. & Dick, R. P., 2001. Seasonal, soil type, and alternative management influences on microbial communities of vegetable cropping systems. Biol Fertil Soils 34, 397–410. https://doi.org/10.1007/s00374-001-0423-7

Schutter, M. E., & Dick, R. P., 2002. Microbial Community Profiles and Activities among Aggregates of Winter Fallow and Cover-Cropped Soil. Soil Science Society of America Journal, 66(1), 142–153. https://doi.org/10.2136/sssaj2002.1420

Shukla, M. K., Lal, R., & Ebinger, M., 2004. Principle Component Analysis for Predicting Corn Biomass And Grain Yields. Soil Science, 169(3), 215–224. https://doi.org/10.1097/01.ss.0000122521.03492.eb

Simmons, B. L., & Coleman, D. C., 2008. Microbial community response to transition from conventional to conservation tillage in cotton fields. Applied Soil Ecology, 40(3), 518–528. https://doi.org/10.1016/j.apsoil.2008.08.003

Singh M. & Siler T., 2022. Soybean planting considerations for maximum profits, MSU Cropping Systems Agronomy website, https://www.canr.msu.edu/news/soybean-planting-and-time-management-considerations

Six, J., Paustian, K., Elliott, E. T., & Combrink, C., 2000. Soil Structure and Organic Matter. Soil Science Society of America Journal, 64(2), 681. https://doi.org/10.2136/sssaj2000.642681x

Soil Survey Staff., 2019. Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online. Accessed [Jan/2019-2023]

Soon Y. K., Abboud S., 1991. A comparison of some methods for soil organic carbon determination, Communications in Soil Science and Plant Analysis, 22:9-10, 943-954, DOI: 10.1080/00103629109368465

Sprunger, C. D., Martin, T., & Mann, M., 2020. Systems with greater perenniality and crop diversity enhance soil biological health. Agricultural & Environmental Letters, 5(1). https://doi.org/10.1002/ael2.20030

Stromberger, M., Shah, Z., & Westfall, D., 2007. Soil microbial communities of no-till dryland agroecosystems across an evapotranspiration gradient. Applied Soil Ecology, 35(1), 94–106. https://doi.org/10.1016/j.apsoil.2006.05.009

Tabatabai, M.A. 1994. Soil enzymes. In: R.W. Weaver, J.S. Angle, and P.S. Bottomley, editors, Methods of soil analysis: Part 2. Microbiological and biochemical properties. Vol. 5. Madison, WI: SSSA. p. 775–833.

Taguchi, M., Izui, K., Katsuki, H., 1980. Augmentation of cyclopropane fatty acid synthesis under stringent control in Escherichia coli. J. Biochem. 88, 1879–1882.

Tay J.K., Narasimhan B., Hastie T., 2023. "Elastic Net Regularization Paths for All Generalized Linear Models." Journal of Statistical Software, 106(1), 1–31. doi:10.18637/jss.v106.i01.

Tunlid, A., Hoitink, H.A.J., Low, C. and White, D.C., 1989. Characterization of bacteria that suppress Rhizoctonia damping-off in bark compost media by analysis of fatty acid biomarkers. Applied and environmental microbiology, 55(6), pp.1368-1374.

Vallejo, V. E., Dick, R. P., Ziv, A., Terán, W., & Roldán, F., 2011. Soil Microbial Community Composition and Enzyme Activity Responses to an Intensive SilvopastoralSystem of Colombia. Nature Precedings. https://doi.org/10.1038/npre.2011.6265.1

Vestal, J. R., & White, D. C., 1989. Lipid Analysis in Microbial Ecology. BioScience, 39(8), 535–541. https://doi.org/10.2307/1310976

Verstraete, W., & Voets, J. P., 1977. Soil microbial and biochemical characteristics in relation to soil management and fertility. *Soil Biology and Biochemistry*, *9*(4), 253–258. https://doi.org/10.1016/0038-0717(77)90031-1

Veum, K.S., Lorenz, T. & Kremer, R.J., 2019., Phospholipid Fatty Acid Profiles of Soils under Variable Handling and Storage Conditions. Agronomy Journal, 111: 1090-1096. https://doi.org/10.2134/agronj2018.09.0628

Veum, K. S., Acosta-Martinez, V., R. Michael Lehman, Li, C., Cano, A., & Nunes, M. R., 2021. PLFA and EL-FAME Indicators of Microbial Community Composition. ASSA, CSSA and SSSA, 251–288. https://doi.org/10.1002/9780891189831.ch12

Wallis, C. M., Lawrence, D. P., Travadon, R., & Baumgartner, K., 2021. Characterization of grapevine fungal canker pathogens Fatty Acid Methyl Ester (FAME) profiles. Mycologia, 1–11. https://doi.org/10.1080/00275514.2021.1983396

Werner, W., Scherer, H.W. and Olfs, H.W., 1988. Influence of long-term application of sewage sludge and compost from garbage with sewage sludge on soil fertility criteria. *Journal of Agronomy and Crop Science*, *160*(3), pp.173-179.

Wickham H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Wilkinson, S. G., 1988. Gram-negative bacteria, p. 299-457. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Yaroschevich, I.V., 1966. Effect of fifty years' application of fertilizers in a rotation on the biological activity of a chernozem. Agrokhimiya 6:14-19.

Yeates, G. W., Orchard, V. A., Speir, T. W., Hunt, J. A., & Hermans, C., 1994. Impact of pasture contamination by copper, chromium, arsenic timber preservative on soil biological activity. Biology and Fertility of Soils, 18(3), 200–208. https://doi.org/10.1007/bf00647667

Zelles, L., Bai, Q. Y., Ma, R. X., Rackwitz, R., Winter, K., & Beese, F., 1994. Microbial biomass, metabolic activity and nutritional status determined from fatty acid patterns and poly-hydroxybutyrate in agriculturally-managed soils. *Soil Biology and Biochemistry*, *26*(4), 439–446. https://doi.org/10.1016/0038-0717(94)90175-9

Zelles, L., 1997. Phospholipid fatty acid profiles in selected members of soil microbial communities. Chemosphere, 35(1-2), 275–294. https://doi.org/10.1016/s0045-6535(97)00155-0

Zelles, L., 1999. Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. Biology and Fertility of Soils, 29(2), 111–129. https://doi.org/10.1007/s003740050533

Table 3.1. Differentiation Power Scoring Table based on Tukey's post hoc testing (P<0.05) for individual long-term field sites.

2 - Treatments (Waterman Farm)	**Post-Hoc outcomes**	a	b		
	Scoring	0	1		
3 - Treatments (Hoytville, Wooster)	**Post-Hoc outcomes**	a	b	c	
	Scoring	0	0.5	1	
4 - Treatments (KBS - Michigan)	**Post-Hoc outcomes**	a	b	c	d
	Scoring	0	0.5	0.75	1

Table 3.2. Table of multivariate soybean yield prediction models with various combinations of variables that are separated into soil property, land management and large variable combination groups (n=153). R^2 and adjusted R^2 were determined for general multivariate linear models and the pseudo R^2 was determine based on a generalized multivariate linear model using the McFadden formula.

	Var. Group	Farm. Type	Soil Texture	SOC + TN	pH Set	Enzy. Act.	Enzy. Ratios	EL-FAME	FAME Ratios	Envi. Factors	Tillage	Plant. Date	Crop Rot.	CC	Res. Cover.	Manure	Pest.	General Linear Model		GLM	Total var. #
	# of var.	1	3	2	6	7	6	40	6	3	5	3	2	4	6	6	5	R^2	Adj. R^2	Pseudo R^2	
	Run #																				
Soil related properties	1	x	x															0.390	0.373	0.390	4
	2	x	x	x														0.426	0.402	0.426	6
	3	x	x	x	x													0.455	0.413	0.455	12
	4	x	x	x	x	x												0.483	0.430	0.483	19
	5	x	x	x	x	x	x											0.545	0.480	0.545	25
	6	x	x	x	x			x										0.648	0.491	0.657	52
	7	x	x	x	x			x	x									0.673	0.499	0.682	58
	8	x	x	x	x	x		x										0.658	0.491	0.666	59
	9	x				x												0.409	0.393	0.409	8
	10	x	x			x												0.437	0.410	0.437	11
	11	x				x	x											0.455	0.421	0.455	14
	12	x	x			x	x											0.488	0.444	0.488	17
	13	x						x										0.606	0.479	0.610	41
	14	x	x					x										0.612	0.473	0.616	44
	15	x						x	x									0.635	0.491	0.638	47
	16	x	x					x	x									0.642	0.486	0.645	50
	17	x				x		x										0.617	0.480	0.620	48
	18	x				x	x	x										0.637	0.484	0.640	54

continues

	19	x				x		x	x									0.645	0.490	0.647	54
	20	x				x	x	x	x									0.678	0.515	0.679	60
	21	x	x		x	x	x	x	x									0.701	0.512	0.702	69
	22	x	x	x		x	x	x	x									0.732	0.576	0.737	65
	23	x	x	x	x	x	x	x	x									0.740	0.566	0.744	71
	24	x	x	x	x	x	x	x	x	x								**0.762**	**0.590**	**0.766**	74
	25		x	x	x	x	x	x	x	x								0.603	0.322	0.606	73
	26		x	x	x	x	x	x	x									0.593	0.327	0.595	70

	Var. Group	Farm. Type	Soil Texture	SOC + TN	pH Set	Enzy. Activ.	Enzy. Ratios	EL-FAME	FAME Ratios	Envir. Factors	Tillage	Plant. Date	Crop Rot.	CC	Res. Cover.	Manure	Pest.	R^2	Adj. R^2	Pseudo R^2	Total var. #
Land Management and Environmental Factors	27										x	x	x	x	x	x	x	0.376	0.259	0.376	31
	28		x							x	x	x	x	x	x	x	x	0.465	0.333	0.465	37
	29	x									x	x	x	x	x	x	x	0.545	0.455	0.545	32
	30	x	x								x	x	x	x	x	x	x	0.610	0.522	0.610	35
	31	x	x							x	x	x	x	x	x	x	x	**0.622**	**0.525**	**0.622**	38
	32	x	x							x		x	x	x	x	x	x	0.590	0.502	0.590	33
	33	x	x							x			x	x	x	x	x	0.590	0.510	0.590	30
	34	x	x							x				x	x	x	x	0.572	0.492	0.572	28
	35	x	x							x					x	x	x	0.571	0.502	0.571	24
	36	x	x							x						x	x	0.543	0.489	0.543	18
	37	x	x							x							x	0.435	0.391	0.435	12
	38	x	x							x								0.419	0.391	0.419	7

continues

	Var. Group	Farm. Type	Soil Texture	SOC + TN	pH Set	Enzy. Activ.	Enzy. Ratios	EL-FAME	FAME Ratios	Envir. Factors	Tillage	Plant. Date	Crop Rot.	CC	Res. Cover.	Manure	Pest.	R^2	Adj. R^2	Pseudo R^2	Total var. #
Large Variable Combinations	39	x	x	x	x					x	x	x	x	x	x	x	x	0.652	0.536	0.652	46
	40	x	x	x	x	x				x	x	x	x	x	x	x	x	0.696	0.584	0.696	53
	41	x	x	x	x	x	x			x	x	x	x	x	x	x	x	0.760	0.656	0.760	59
	42	x	x	x	x			x		x	x	x	x	x	x	x	x	0.782	0.574	0.796	86
	43	x	x	x	x			x	x	x	x	x	x	x	x	x	x	0.789	0.555	0.804	92
	44	x	x	x	x	x		x		x	x	x	x	x	x	x	x	0.819	0.634	0.839	93
	45	x	x	x		x	x	x		x	x	x	x	x	x	x	x	0.849	0.694	0.861	93
	46	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	0.859	0.689	0.872	99
	47	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	**0.860**	**0.695**	**0.870**	99
	48	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0.870	0.690	0.882	105
	49	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	0.854	0.695	0.866	102
	50	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	0.846	0.679	0.859	102
	51	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	0.838	0.658	0.843	103
	52	x	x			x	x	x	x	x	x	x	x	x	x	x	x	0.829	0.663	0.835	97

Table 3.3. Example of how the coef. ratios ($\beta_{p\text{-}f}$) were calculated for individual enzyme activity variables. In this example the regression coefficients (β_p) were extracted from the amended multivariate regression model.

		Variable (x_i)	*regress. coefficient (β_p)*
Enzyme	**Enzyme Activity**	GLU	0
		NAG	6524.352
		AS	-1668.738
		GLUAS	-5558.141
		GLUNAG	0
		NAGAS	0
		GLUNAGAS	0
	Enzyme Ratios	GLU / AS	1698.396
		GLU / NAG	-2229.598
		NAG / AS	0
		(GLU + AS) / NAG	6396.022
		(GLU + NAG) / AS	0
		(AS + NAG) / GLU	2523.709

Positive **Coefficient Ratio**

Variable (x_i)	*(β_p)*	coef. ratio ($\beta_{p\text{-}f}$)
NAG	6524.352	0.3806
GLU / AS	1698.396	0.0991
(GLU + AS) / NAG	6396.022	0.3731
(AS + NAG) / GLU	2523.709	0.1472
Sum	**17142.479**	

Negative **Coefficient Ratio**

Variable (x_i)	*(β_p)*	coef. ratio ($\beta_{p\text{-}f}$)
AS	-1668.738	0.1765
GLUAS	-5558.141	0.5878
GLU / NAG	-2229.598	0.2358
Sum	**-9456.477**	

Table 3.4. Example of how the mean weighted positive, negative, and fitted biochemical Soil Health Scores are calculated. In this example the measurements for the three replication plots at the LTFS in Wooster, Ohio were used. The individual normalized explanatory variables (x_i) measurements are multiplied with the corresponding coef. ratio ($\beta_{p\text{-}f}$) to determine individual SHS for specific variables. Afterwards the individual SHS results are summed up to determine the positive or negative SHS for each replication. To determine the weighted fitted SHS a new positive and negative coef. ratio is calculated and then multiplied with the corresponding summed SHS result.

Variable	NAG			GLU / AS			(GLU + AS) / NAG			(AS + NAG) / GLU			*Positive*
Rep.	x_i	$\beta_{p\text{-}f}$	SHS_i	x_i	$\beta_{p\text{-}f}$	SHS_i	x_i	$\beta_{p\text{-}f}$	SHS_i	x_i	$\beta_{p\text{-}f}$	SHS_i	**$\sum(SHS_i)$**
1	0.362	0.381	0.138	0.644	0.099	0.064	0.446	0.373	0.167	0.236	0.147	0.035	*0.403*
2	0.370	0.381	0.141	0.596	0.099	0.059	0.454	0.373	0.169	0.251	0.147	0.037	*0.406*
3	0.457	0.381	0.174	0.513	0.099	0.051	0.374	0.373	0.140	0.299	0.147	0.044	*0.408*
												Mean:	**0.406**

Variable	AS			GLU + AS			GLU / NAG			*Negative*
Rep.	x_i	$\beta_{p\text{-}f}$	SHS_i	x_i	$\beta_{p\text{-}f}$	SHS_i	x_i	$\beta_{p\text{-}f}$	SHS_i	**$\sum(SHS_i)$**
1	0.218	0.176	0.039	0.350	0.588	0.206	0.472	0.236	0.111	*0.356*
2	0.238	0.176	0.042	0.364	0.588	0.214	0.467	0.236	0.110	*0.366*
3	0.265	0.176	0.047	0.370	0.588	0.218	0.362	0.236	0.085	*0.350*
									Mean:	**0.357**

CS_PT	positive SHS			negative SHS			*Fitted*
Rep.	SHS +	$\beta_{p\text{-}f}$	SHS_i	SHS -	$\beta_{p\text{-}f}$	SHS_i	**$\sum(SHS_i)$**
1	0.403	0.644	0.260	0.356	0.356	0.126	*0.386*
2	0.406	0.644	0.262	0.366	0.356	0.130	*0.392*
3	0.408	0.644	0.263	0.350	0.356	0.124	*0.388*
						Mean:	**0.389**

Fitted Coefficient Ratio

Variable (x_i)	$\sum \beta_p$	coef. ratio ($\beta_{p\text{-}f}$)
SHS +	17142.479	*0.644*
|SHS -|	9456.477	*0.356*
Sum	**26598.956**	

Table 3.5. Table of multivariate soybean yield prediction models developed from various variable combinations that include Soil Health Scores, land management factors, environmental factors, soil properties, and the binary variable for org. vs conv. farming.

		Variable Group	Farm. Type	Soil Texture	SOC + TN	pH Set	*Enzyme - SHS*		*EL-FAME - SHS*		*Enzyme + EL-FAME - SHS*		Envi. Factors	All Land Manag. Factors	General Linear Model		GLM	Total var. #
							pos. + neg.	fitted	pos.+ neg.	fitted	pos.+ neg.	fitted						
		# of var.	1	3	2	6	2	1	2	1	2	1	3	31	R^2	Adj. R^2	Pseudo R^2	
		Run #																
Simplified Models that included Soil Health Score variables	Enzyme	1	x	x	x	x		x					x	x	0.653	0.533	0.653	47
		2	x	x	x	x	x						x	x	0.752	0.663	0.752	48
		3	x	x	x	x	x							x	0.742	0.659	0.742	45
		4	x		x	x	x						x	x	0.744	0.661	0.744	45
		5	x	x		x	x						x	x	0.729	0.639	0.729	46
		6		x	x	x	x						x	x	0.595	0.455	0.595	47
		7	x	x	x	x	x						x		0.523	0.467	0.523	17
	EL-FAME	8	x	x	x	x				x			x	x	0.655	0.536	0.655	47
		9	x	x	x	x			x				x	x	0.726	0.629	0.726	48
		10	x	x	x	x			x					x	0.713	0.620	0.713	45
		11	x		x	x			x				x	x	0.701	0.605	0.701	45
		12	x	x		x			x				x	x	0.685	0.580	0.685	46
		13		x	x	x			x				x	x	0.568	0.418	0.568	47
		14	x	x	x	x			x				x		0.543	0.490	0.543	17
	Enzyme + EL-FAME	15	x	x	x	x						x	x	x	0.655	0.535	0.655	47
		16	x	x	x	x					x		x	x	**0.841**	**0.784**	**0.841**	48
		17	x	x	x	x					x			x	0.828	0.773	0.828	45
		18	x		x	x					x		x	x	0.821	0.764	0.821	45
		19	x	x		x					x		x	x	0.792	0.723	0.792	46
		20		x	x	x					x		x	x	0.650	0.529	0.650	47
		21	x	x	x	x					x		x		0.584	0.535	0.584	17
		22	x	x		x	x							x	0.711	0.640	0.711	43
		23	x	x		x			x					x	0.673	0.592	0.673	43
		24	x	x		x					x			x	0.773	0.717	0.773	43

Table 3.6. Ester-linked fatty acid methyl ester (EL-FAME) results [a)] of four different long-term field sites with different treatment practices taken at a soil depth of 15 cm (KBS-LTER site in Michigan, Triplett-Van Doren Sites in Wooster and Hoytville, and the East Straw Mulch Experiment located at the Waterman Farm in Columbus, Ohio).

Site	Treatment	n	Total FAMEs	Gram-positive (+) bacteria	Gram-negative (-) bacteria	Actinomycetes	Arbuscular mycorrhizal fungi (AMF)	Protozoa	Eukaryotes	16:0 iso [Gram (+)]
			nmol g^{-1} soil							
KBS-LTER (Michigan)	CSW-MT-Conv [b)]	6	98.8 (± 8.0) b	11.2 (± 0.9) c	14.1 (± 1.6) c	6.32 (± 0.66) c	2.58 (± 0.52) c	0.956 (± 0.13) b	6.36 (± 0.71) c	2.49 (± 0.22) c
	CSW-NT-Conv	6	102.3 (± 13.1) b	12.5 (± 1.6) c	16.0 (± 2.2) c	8.19 (± 1.20) b	3.42 (± 0.42) bc	1.156 (± 0.21) ab	7.01 (± 0.84) bc	2.88 (± 0.39) bc
	CSW-MT-CC-Conv(-)	6	131.7 (± 9.1) a	15.8 (± 1.2) b	19.6 (± 0.9) b	8.69 (± 0.47) b	4.61 (± 1.13) ab	1.419 (± 0.19) a	8.43 (± 0.70) ab	3.46 (± 0.34) ab
	CSW-MT-CC-Org	6	146.7 (± 16.1) a	18.7 (± 2.2) a	22.9 (± 2.9) a	10.66 (± 1.23) a	5.01 (± 0.99) a	1.176 (± 0.19) ab	9.74 (± 1.54) a	3.85 (± 0.51) a
Hoytville (Ohio)	CS_PT	3	150.2 (± 14.1) b	21.3 (± 2.0) b	23.4 (± 2.8) a	13.2 (± 1.33) b	7.73 (± 1.10) a	1.062 (± 0.30) a	6.86 (± 0.39) b	3.98 (± 0.23) c
	CS_MT	3	182.8 (± 8.1) a	26.8 (± 0.9) a	25.5 (± 1.4) a	15.5 (± 0.79) ab	9.36 (± 1.43) a	1.001 (± 0.15) a	8.74 (± 0.34) a	5.29 (± 0.17) b
	CS_NT	3	186.5 (± 8.0) a	30.5 (± 1.8) a	28.6 (± 3.1) a	17.0 (± 1.19) a	7.71 (± 0.96) a	0.906 (± 0.12) a	9.44 (± 0.47) a	6.57 (± 0.36) a
Wooster (Ohio)	CS_PT	3	144.1 (± 20.4) a	17.7 (± 1.5) ab	23.1 (± 4.4) a	9.95 (± 0.94) a	5.48 (± 0.97) a	1.176 (± 0.10) a	9.41 (± 1.34) ab	3.60 (± 0.30) ab
	CS_MT	3	186.6 (± 51.3) a	20.3 (± 2.8) a	23.0 (± 3.9) a	11.5 (± 2.01) a	6.31 (± 0.89) a	1.333 (± 0.09) a	10.88 (± 1.47) a	3.93 (± 0.34) a
	CS_NT	3	114.3 (± 22.6) a	15.2 (± 1.4) b	18.3 (± 4.2) a	8.61 (± 0.92) a	6.14 (± 2.29) a	1.187 (± 0.36) a	7.47 (± 1.16) b	3.01 (± 0.27) b
Waterman Farm (Ohio)	M0F0	3	57.5 (± 9.2) b	8.1 (± 0.84) b	8.97 (± 1.3) b	5.46 (± 0.87) b	1.17 (± 0.34) b	0.363 (± 0.13) b	3.63 (± 0.96) b	1.51 (± 0.27) b
	M16F0	3	96.9 (± 13.1) a	13.4 (± 1.1) a	15.0 (± 2.9) a	7.76 (± 0.28) a	3.29 (± 0.22) a	0.701 (± 0.17) a	6.11 (± 1.05) a	2.18 (± 0.17) a
Diff. Power *(max = 4)*			***2***	***2.75***	***1.75***	***2.25***	***1.75***	***1.5***	***2.75***	***3.25***

a) Data are means and standard deviation values across the four long-term field sites

b) CSW-MT-Conv…Conventional; CSW-NT-Conv…No-Till; CSW-MT-CC-Conv(-)…Conventional with Reduced Input (30% less N) with Cover Crops; CSW-MT-CC-Org …Biologically Based with Cover Crops

c) CS…Corn/Soy crop rotation; PT…Plow Till (Moldboard Plow); MT…Minimal tillage (Chisel Till); NT…No-Till

d) M0F0…No-Till + No Mulch + No Fertilizer; M16F0…No-Till + 16Mg/ha Mulch + No Fertilizer

e) Sum of a15:0, i15:0, i16:0, a17:0, and i17:0

f) Sum of cy17:0, cy19:0, 16:1ω7c, 17:1ω8c, 18:1ω7c

g) Sum of 10Me16:0, 10Me17:0, 10Me18:0, 10Me19:1ω7c

h) 16:1ω5c

I) Sum of 20:3ω6c, and 20:4ω6c

j) Sum of 21:0, 22:0, 23:0 and 24:0

k) Values within a site and column followed by the same letter(s) are not significantly different according to Tukey test (P< 0.05)

Table 3.7. Ester-linked fatty acid methyl ester (EL-FAME) - Ratio results [a] from four long-term field sites with different treatment practices taken at a soil depth of 15 cm (KBS-LTER site in Michigan, Triplett-Van Doren Sites in Wooster and Hoytville, and the East Straw Mulch Experiment located at the Waterman Farm in Columbus, Ohio).

Site	Treatment	n	Bacteria [e]/ Total FAMEs	Fungi [f]/ Total FAMEs	Fungi [f]/ Bacteria [e]	AMF [g]/ Bacteria [e]	SAT [h]/ MONOSAT [i]	Gram (+) [j]/ Gram (-) [k]	cy17:0 / Precursor [l]	cy19:0 / Precursor [m]
			(-)							
KBS-LTER	CSW-MT-Conv [b]	6	0.330 (± 0.01) b [n]	0.281 (± 0.03) a	0.853 (± 0.11) a	0.079 (± 0.01) b	0.724 (± 0.06) a	0.798 (± 0.06) a	0.252 (± 0.02) a	0.602 (± 0.07) a
	CSW-NT-Conv	6	0.371 (± 0.01) a	0.223 (± 0.01) b	0.602 (± 0.04) c	0.092 (± 0.01) ab	0.698 (± 0.05) a	0.786 (± 0.07) a	0.248 (± 0.04) a	0.610 (± 0.05) a
	CSW-MT-CC-Conv(-)	6	0.345 (± 0.01) b	0.248 (± 0.01) b	0.721 (± 0.04) b	0.101 (± 0.02) a	0.708 (± 0.05) a	0.810 (± 0.04) a	0.260 (± 0.03) a	0.463 (± 0.03) b
	CSW-MT-CC-Org	6	0.366 (± 0.01) a	0.236 (± 0.01) b	0.646 (± 0.06) bc	0.093 (± 0.01) ab	0.725 (± 0.03) a	0.820 (± 0.06) a	0.254 (± 0.01) a	0.517 (± 0.04) b
Hoytville (Ohio)	CS_PT [c]	3	0.394 (± 0.01) ab	0.210 (± 0.01) a	0.534 (± 0.01) a	0.131 (± 0.01) a	0.724 (± 0.13) a	0.913 (± 0.02) b	0.309 (± 0.04) a	0.658 (± 0.18) b
	CS_MT	3	0.381 (± 0.00) b	0.209 (± 0.01) a	0.548 (± 0.02) a	0.135 (± 0.03) a	0.804 (± 0.02) a	1.052 (± 0.02) a	0.340 (± 0.03) a	0.738 (± 0.09) b
	CS_NT	3	0.419 (± 0.02) a	0.172 (± 0.00) b	0.411 (± 0.02) b	0.099 (± 0.01) a	0.853 (± 0.04) a	1.073 (± 0.06) a	0.284 (± 0.04) a	1.039 (± 0.02) a
Wooster (Ohio)	CS_PT [c]	3	0.364 (± 0.03) a	0.236 (± 0.06) a	0.660 (± 0.23) a	0.105 (± 0.01) a	0.667 (± 0.06) a	0.791 (± 0.19) a	0.220 (± 0.07) a	0.434 (± 0.05) a
	CS_MT	3	0.319 (± 0.10) a	0.293 (± 0.15) a	1.112 (± 0.96) a	0.113 (± 0.01) a	0.689 (± 0.08) a	0.886 (± 0.03) a	0.227 (± 0.01) a	0.473 (± 0.02) a
	CS_NT	3	0.379 (± 0.02) a	0.216 (± 0.01) a	0.571 (± 0.01) a	0.139 (± 0.03) a	0.696 (± 0.05) a	0.853 (± 0.12) a	0.278 (± 0.05) a	0.497 (± 0.10) a
Waterman Farm (Ohio)	M0F0 [d]	3	0.423 (± 0.03) a	0.220 (± 0.05) a	0.524 (± 0.14) a	0.05 (± 0.01) b	0.695 (± 0.10) a	0.911 (± 0.11) a	0.214 (± 0.17) a	0.712 (± 0.09) a
	M16F0	3	0.388 (± 0.01) a	0.194 (± 0.02) a	0.502 (± 0.07) a	0.089 (± 0.01) a	0.768 (± 0.07) a	0.912 (± 0.10) a	0.274 (± 0.04) a	0.467 (± 0.04) b
Diff. Power *(max = 4)*			***1***	***1***	***1.25***	***1.5***	***0***	***0.5***	***0***	***2***

a) Data are means and standard deviation values across the four long-term field sites

b) CSW-MT-Conv…Conventional; CSW-NT-Conv…No-Till; CSW-MT-CC-Conv(-)…Conventional with Reduced Input (30% less N) with Cover Crops; CSW-MT-CC-Org …Biologically Based with Cover Crops

c) CS…Corn/Soy crop rotation; PT…Plow Till (Moldboard Plow); MT…Minimal tillage (Chisel Till); NT…No-Till

d) M0F0…No-Till + No Mulch + No Fertilizer; M16F0…No-Till + 16Mg/ha Mulch + No Fertilizer

e) Sum of 15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c

f) Sum of 18:1ω9c, 18:2ω6c, and 16:1ω5c

g) 16:1ω5c

h) Sum of 14:0, 15:0, 16:0, 17:0, and 18:0

i) Sum of 16:1ω5c, 16:1ω7c, 17:1ω8c, 18:1ω7c, 18:1ω9c, cy17:0, and cy19:0

j) Sum of a15:0, i15:0, i16:0, a17:0, and i17:0

k) Sum of cy17:0, cy19:0, 16:1ω7c, 17:1ω8c, and 18:1ω7c

l) 16:1ω7c m) 18:1ω7c

n) Individual long-term field site means within a column followed by the same letter(s) are not significantly different according to Tukey test (P< 0.05).

Table 3.8. Soil enzyme activity results [a)] from four long-term field sites with different treatment practices taken at a soil depth of 15 cm (KBS-LTER site in Michigan, Triplett-Van Doren Sites in Wooster and Hoytville, and the East Straw Mulch Experiment located at the Waterman Farm in Columbus, Ohio).

Site	Treatment	n	β-Glucosidase activity (β-GLU)	Arylsulfatase activity (AS)	N-acetyl-β-glucosaminidase activity (NAG)	β-GLU + AS	β-GLU + NAG	AS + NAG	β-GLU + NAG + AS
			mg PNP g^{-1} dry soil h^{-1}						
KBS-LTER (Michigan)	CSW-MT-Conv [b)]	6	239.2 (± 33.5) b	112.8 (± 47.8) b	72.6 (± 8.0) b	352.0 (± 78.6) c	311.8 (± 33.3) b	185.4 (± 44.7) b	424.6 (± 76.2) c
	CSW-NT-Conv	6	277.5 (± 33.1) b	210.1 (± 38.4) a	87.0 (± 10.4) b	487.6 (± 68.9) b	364.5 (± 41.7) b	297 (± 46.9) a	574.6 (± 77.5) b
	CSW-MT-CC-Conv(-)	6	379.5 (± 38.6) a	217.4 (± 24.9) a	113.0 (± 9.8) a	596.9 (± 61.8) ab	492.6 (± 47.3) a	330.4 (± 33.2) a	709.9 (± 70.3) a
	CSW-MT-CC-Org	6	429.5 (± 49.6) a	246.4 (± 49.8) a	117.5 (± 7.9) a	675.9 (± 95.2) a	546.9 (± 53.8) a	363.9 (± 56.5) a	793.3 (± 100.6) a
Hoytville (Ohio)	CS_PT	3	458.4 (± 38.5) a	446.8 (± 35.5) c	100.6 (± 6.8) a	905.3 (± 29.4) c	559.1 (± 44.7) a	547.5 (± 29.3) c	1005.9 (± 30.5) c
	CS_MT	3	546.9 (± 54.3) a	579.6 (± 41.7) b	119.0 (± 26.4) a	1126.6 (± 95.8) b	666 (± 79.8) a	698.7 (± 66.3) b	1245.6 (± 120.6) b
	CS_NT	3	550.9 (± 27.9) a	806.2 (± 44.2) a	111.1 (± 6.2) a	1357.1 (± 66.3) a	662 (± 32.9) a	917.3 (± 44.9) a	1468.2 (± 68.7) a
Wooster (Ohio)	CS_PT	3	552.3 (± 12.3) a	345.2 (± 33.7) a	122.5 (± 16.2) a	897.6 (± 25.4) a	674.9 (± 5.9) a	467.8 (± 49.2) a	1020.1 (± 39.3) a
	CS_MT	3	489.8 (± 37.5) a	329.6 (± 49.7) a	116.1 (± 4.8) a	819.3 (± 86.7) a	605.8 (± 33.5) a	445.6 (± 45.2) a	935.4 (± 82.3) a
	CS_NT	3	431.4 (± 90.8) a	408.9 (± 79.2) a	125.4 (± 8.7) a	840.4 (± 158.9) a	556.8 (± 96.8) a	534.3 (± 87.8) a	965.7 (± 166.5) a
Waterman Farm (Ohio)	M0F0	3	113.7 (± 67.8) b	198.1 (± 92.1) b	33.3 (± 6.5) b	311.7 (± 158.9) b	147.0 (± 73.1) b	231.4 (± 96.5) b	345 (± 163.7) b
	M16F0	3	232.6 (± 29.9) a	440.6 (± 55.2) a	82.0 (± 20.4) a	673.3 (± 78.9) a	314.6 (± 44.5) a	522.6 (± 75.4) a	755.2 (± 97.6) a
Diff. Power *(max = 4)*			***1.5***	***2.5***	***1.5***	***2.75***	***1.5***	***2.5***	***2.75***

a) Data as means including the computed standard deviation.
b) CSW-MT-Conv…Conventional; CSW-NT-Conv…No-Till; CSW-MT-CC-Conv(-)…Conventional with Reduced Input (30% less N) with Cover Crops; CSW-MT-CC-Org …Biologically Based with Cover Crops
c) CS…Corn/Soy crop rotation; PT…Plow Till (Moldboard Plow); MT…Minimal tillage (Chisel Till); NT…No-Till
d) M0F0…No-Till + No Mulch + No Fertilizer; M16F0…No-Till + 16Mg/ha Mulch + No Fertilizer
e) p-nitrophenol
f) Individual long-term field site means within a column followed by the same letter(s) are not significantly different according to Tukey test (P< 0.05)

Table 3.9. Soil enzyme activity ratio results [a)] from four long-term field sites with different treatment practices taken at a soil depth of 15 cm (KBS-LTER site in Michigan, Triplett-Van Doren Sites in Wooster and Hoytville, and the East Straw Mulch Experiment located at the Waterman Farm in Columbus, Ohio).

Site	Treatment	n	β-GLU / NAG	β-GLU / AS	NAG / AS	(β-GLU + NAG) / AS	(β-GLU + AS) / NAG	(NAG + AS) / β-GLU
			(-)					
KBS-LTER (Michigan)	CSW-MT-Conv [b)]	6	3.335 (± 0.67) a [f)]	2.271 (± 0.46) a	0.716 (± 0.22) a	2.987 (± 0.67) a	4.94 (± 1.47) a	0.769 (± 0.09) b
	CSW-NT-Conv	6	3.200 (± 0.25) a	1.337 (± 0.14) c	0.419 (± 0.04) b	1.757 (± 0.17) b	5.61 (± 0.44) a	1.068 (± 0.09) a
	CSW-MT-CC-Conv(-)	6	3.358 (± 0.17) a	1.750 (± 0.10) bc	0.522 (± 0.04) b	2.272 (± 0.13) b	5.28 (± 0.27) a	0.871 (± 0.04) b
	CSW-MT-CC-Org	6	3.658 (± 0.35) a	1.772 (± 0.21) b	0.487 (± 0.07) b	2.259 (± 0.26) b	5.74 (± 0.59) a	0.847 (± 0.08) b
Hoytville (Ohio)	CS_PT [c)]	3	4.554 (± 0.17) a	1.033 (± 0.16) a	0.227 (± 0.03) a	1.260 (± 0.19) a	9.02 (± 0.66) b	1.202 (± 0.14) b
	CS_MT	3	4.676 (± 0.56) a	0.942 (± 0.03) a	0.204 (± 0.03) ab	1.146 (± 0.06) a	9.65 (± 1.29) b	1.278 (± 0.01) b
	CS_NT	3	4.962 (± 0.19) a	0.684 (± 0.03) b	0.138 (± 0.01) b	0.822 (± 0.04) b	12.23 (± 0.74) a	1.666 (± 0.05) a
Wooster (Ohio)	CS_PT [c)]	3	4.565 (± 0.65) a	1.612 (± 0.18) a	0.354 (± 0.02) a	1.966 (± 0.17) a	7.39 (± 0.77) a	0.848 (± 0.11) b
	CS_MT	3	4.232 (± 0.49) a	1.497 (± 0.11) a	0.359 (± 0.06) a	1.855 (± 0.17) a	7.09 (± 1.04) a	0.909 (± 0.03) b
	CS_NT	3	3.430 (± 0.57) a	1.062 (± 0.17) b	0.312 (± 0.04) a	1.374 (± 0.18) b	6.67 (± 0.86) a	1.255 (± 0.20) a
Waterman Farm (Ohio)	M0F0 [d)]	3	3.296 (± 0.38) a	0.559 (± 0.51) a	0.188 (± 0.04) a	0.746 (± 0.88) a	9.18 (± 0.41) a	2.168 (± 0.54) a
	M16F0	3	2.935 (± 0.17) a	0.530 (± 0.31) a	0.184 (± 0.11) a	0.714 (± 0.44) a	8.43 (± 0.25) a	2.255 (± 0.42) a
Diff. Power *(max =4)*			***0***	***1.75***	***1***	***1.5***	***0.5***	***1.5***

a) Data as means including the computed standard deviation.
b) CSW-MT-Conv…Conventional; CSW-NT-Conv…No-Till; CSW-MT-CC-Conv(-)…Conventional with Reduced Input (30% less N) with Cover Crops; CSW-MT-CC-Org …Biologically Based with Cover Crops
c) CS…Corn/Soy crop rotation; PT…Plow Till (Moldboard Plow); MT…Minimal tillage (Chisel Till); NT…No-Till
d) M0F0…No-Till + No Mulch + No Fertilizer; M16F0…No-Till + 16Mg/ha Mulch + No Fertilizer
e) p-nitrophenol.
f) Individual long-term field site means within a column followed by the same letter(s) are not significantly different according to Tukey test (P< 0.05).

Table 3.10. Soybean Yield, Total Carbon, Soil Org. Carbon, Total Nitrogen, Soil Health Score results [a)] from four long-term field sites with different treatment practices taken at a soil depth of 15 cm (KBS-LTER site in Michigan, Triplett-Van Doren Sites in Wooster and Hoytville, and the East Straw Mulch Experiment located at the Waterman Farm in Columbus, Ohio).

Site	Treatment	n	Soybean Yield	Total Carbon	Soil Organic Carbon	Total Nitrogen	fitted SH-Score (Enzyme)	fitted SH-Score (EL-FAME)	fitted SH-Score (Enzyme + EL-FAME)
			kg ha $^{-1}$	g kg^{-1}			(-)		
KBS-LTER (Michigan)	CSW-MT-Conv [b)]	6	4229 (± 316) a [f)]	9.95 (± 0.77) ab	9.38 (± 0.80) ab	0.88 (± 0.08) b	0.262 (± 0.024) c	0.249 (± 0.025) b	0.253 (± 0.019) c
	CSW-NT-Conv	6	4182 (± 146) a	9.72 (± 0.93) b	9.13 (± 0.90) b	0.90 (± 0.04) b	0.285 (± 0.017) bc	0.271 (± 0.027) b	0.275 (± 0.022) c
	CSW-MT-CC-Conv(-)	6	4097 (± 287) a	11.31 (± 1.22) a	10.75 (± 1.23) a	1.03 (± 0.10) a	0.315 (± 0.015) ab	0.330 (± 0.021) a	0.325 (± 0.016) b
	CSW-MT-CC-Org	6	2006 (± 332) b	10.69 (± 0.95) ab	10.12 (± 0.94) ab	1.00 (± 0.07) ab	0.335 (± 0.021) a	0.380 (± 0.049) a	0.366 (± 0.033) a
Hoytville (Ohio)	CS_PT [c)]	3	3963 (± 1.70) b †	14.47 (± 1.45) b	13.72 (± 1.38) b	1.57 (± 0.11) b	0.396 (± 0.010) c	0.319 (± 0.018) b	0.343 (± 0.015) c
	CS_MT	3	4712 (± 352) a	17.05 (± 0.82) ab	16.27 (± 0.83) ab	1.82 (± 0.12) ab	0.445 (± 0.011) b	0.392 (± 0.015) a	0.408 (± 0.014) b
	CS_NT	3	5006 (± 154) a	17.65 (± 1.34) a	16.91 (± 1.31) a	1.97 (± 0.15) a	0.511 (± 0.017) a	0.422 (± 0.021) a	0.450 (± 0.020) a
Wooster (Ohio)	CS_PT [c)]	3	2786 (± 836) a	9.26 (± 0.74) a	8.64 (± 0.76) a	0.98 (± 0.04) a	0.389 (± 0.003) a	0.356 (± 0.036) ab	0.366 (± 0.026) ab
	CS_MT	3	3220 (± 339) a	10.78 (± 1.15) a	10.16 (± 1.16) a	1.05 (± 0.06) a	0.368 (± 0.022) a	0.412 (± 0.024) a	0.398 (± 0.010) a
	CS_NT	3	2826 (± 95) a	9.48 (± 1.10) a	8.82 (± 1.07) a	1.01 (± 0.07) a	0.369 (± 0.036) a	0.297 (± 0.040) b	0.320 (± 0.039) b
Waterman Farm (Ohio)	M0F0 [d)]	3	1815 (± 491) a	9.69 (± 1.25) b	9.01 (± 1.21) b	0.97 (± 0.04) b	0.291 (± 0.071) a	0.149 (± 0.032) b	0.194 (± 0.033) b
	M16F0	3	1987 (± 1089) a	13.82 (± 1.32) a	13.08 (± 1.35) a	1.15 (± 0.04) a	0.359 (± 0.011) a	0.249 (± 0.038) a	0.284 (± 0.024) a
Diff. Power *(max =4)*			***1***	***2***	***2***	***2***	***1.75***	***2.5***	***3.25***

a) Data as means including the computed standard deviation

b) CSW-MT-Conv…Conventional; CSW-NT-Conv…No-Till; CSW-MT-CC-Conv(-)…Conventional with Reduced Input (30% less N) with Cover Crops; CSW-MT-CC-Org …Biologically Based with Cover Crops

c) CS…Corn/Soy crop rotation; PT…Plow Till (Moldboard Plow); MT…Minimal tillage (Chisel Till); NT…No-Till

d) M0F0…No-Till + No Mulch + No Fertilizer; M16F0…No-Till + 16Mg/ha Mulch + No Fertilizer

e) Individual long-term field site means within a column followed by the same letter(s) are not significantly different according to Tukey test (P< 0.05)

† Only two soybean yield samples were recorded and used

Table 3.11. Soybean yields and Tukey Post Hoc test results (α = 0.05) from four management systems at the Main Cropping System Experiment in Michigan (KBS-LTER). The mean yield is determined for each growing season between 1990 to 2021.

Year	CSW-MT-Conv [a)]	CSW-NT-Conv	CSW-MT-CC-Conv(-)	CSW-MT-CC-Org	ANOVA
	kg /ha				
1990	*2841*	*2973*	*0*	*0*	n.d.
1991	*0*	*0*	*3177*	*3148*	n.d.
1992	*1839*	*2219*	*0*	*0*	n.d.
1994	*3013*	*2862*	*3091*	*3192*	n.d.
1997	*1515 c*	*2044 a*	*1814 ab*	*1651 bc*	*
2000	*2657*	*2896*	*2866*	*2899*	n.d.
2003	*1621 a*	*1855 a*	*1198 b*	*1010 b*	*
2006	*2867 c*	*3602 a*	*3205 b*	*2971 bc*	*
2009	*1967 b*	*2570 a*	*2219 b*	*2165 b*	*
2012	*1302 b*	*1884 a*	*1405 b*	*1215 b*	*
2015	*3537 b*	*3966 a*	*4154 a*	*2530 c*	**
2018	*3101*	*3589*	*3617*	*3315*	n.d.
2021	*4226 a*	*4180 a*	*4094 a*	*2005 b*	*
Mean	**2580 ab**	**2945 a**	**2766 a**	**2295 b**	
Yield Gap [b)]	11.1%	22.1%	17.0%	-	

a) CSW-MT-Conv… Conventional
CSW-NT-Conv… No-Till
CSW-MT-CC-Conv(-)… Conventional with Reduced Input (30% less N) with Cover Crops
CSW-MT-CC-Org … Biologically Based with Cover Crops

b) Percentage difference in soybean yield between T4 and the other three treatment designs

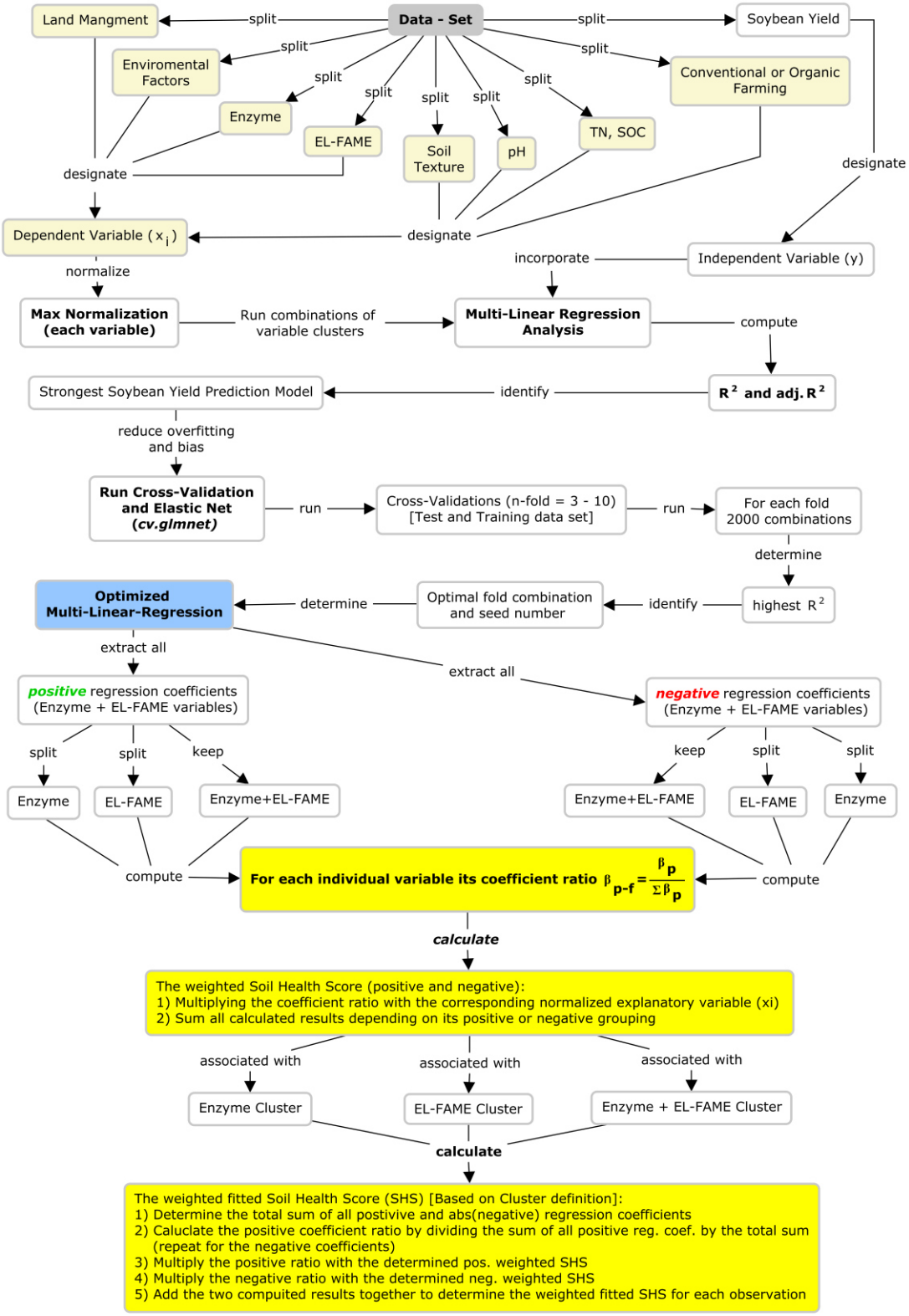


Figure 3.1. Flowchart for the determination of the most reliable multi-linear soybean prediction model including model optimization steps and computation steps to determine the positive, negative, and fitted Soil Health Scores.

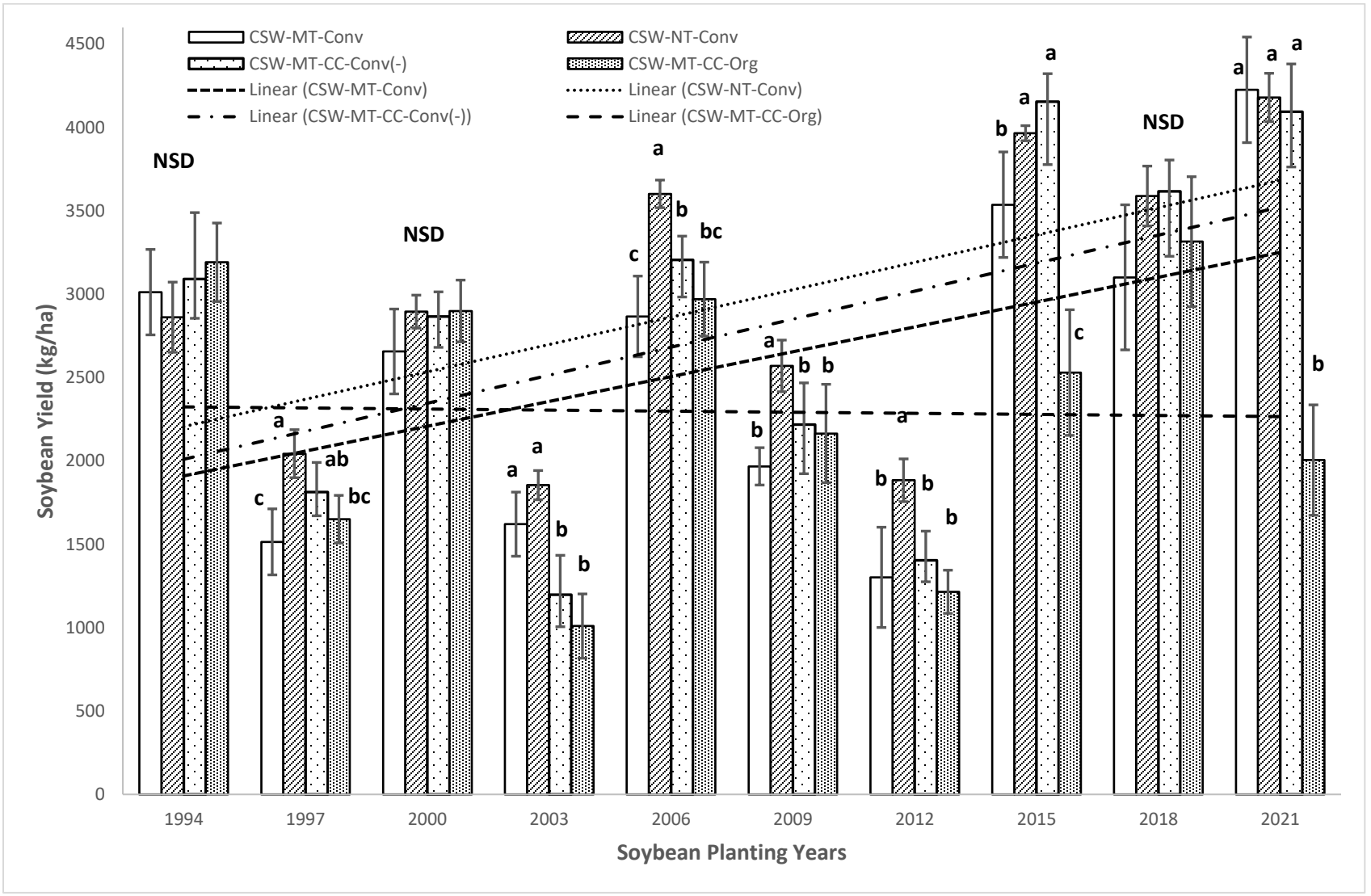


Figure 3.2. Shifts in soybean yields over the past 10 growing seasons at the KBS Long-Term Field Site in Michigan

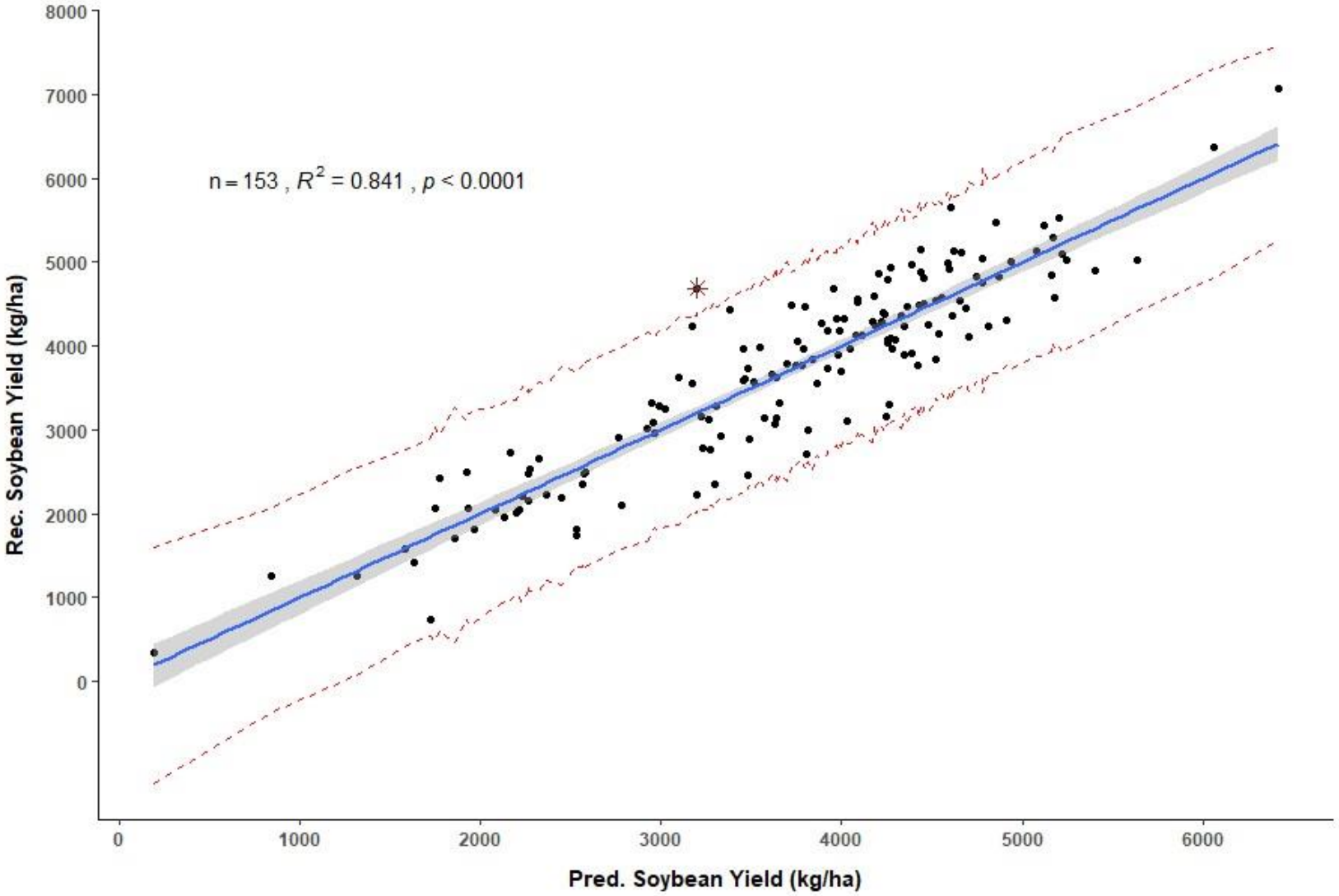


Figure 3.3. Relationship between actual soybean yields and predicted soybean yields of the best model. The gray band around the blue fitted line represents the 95 % confidence interval for the fitted values. The red dashed line represents the 95 % prediction interval.

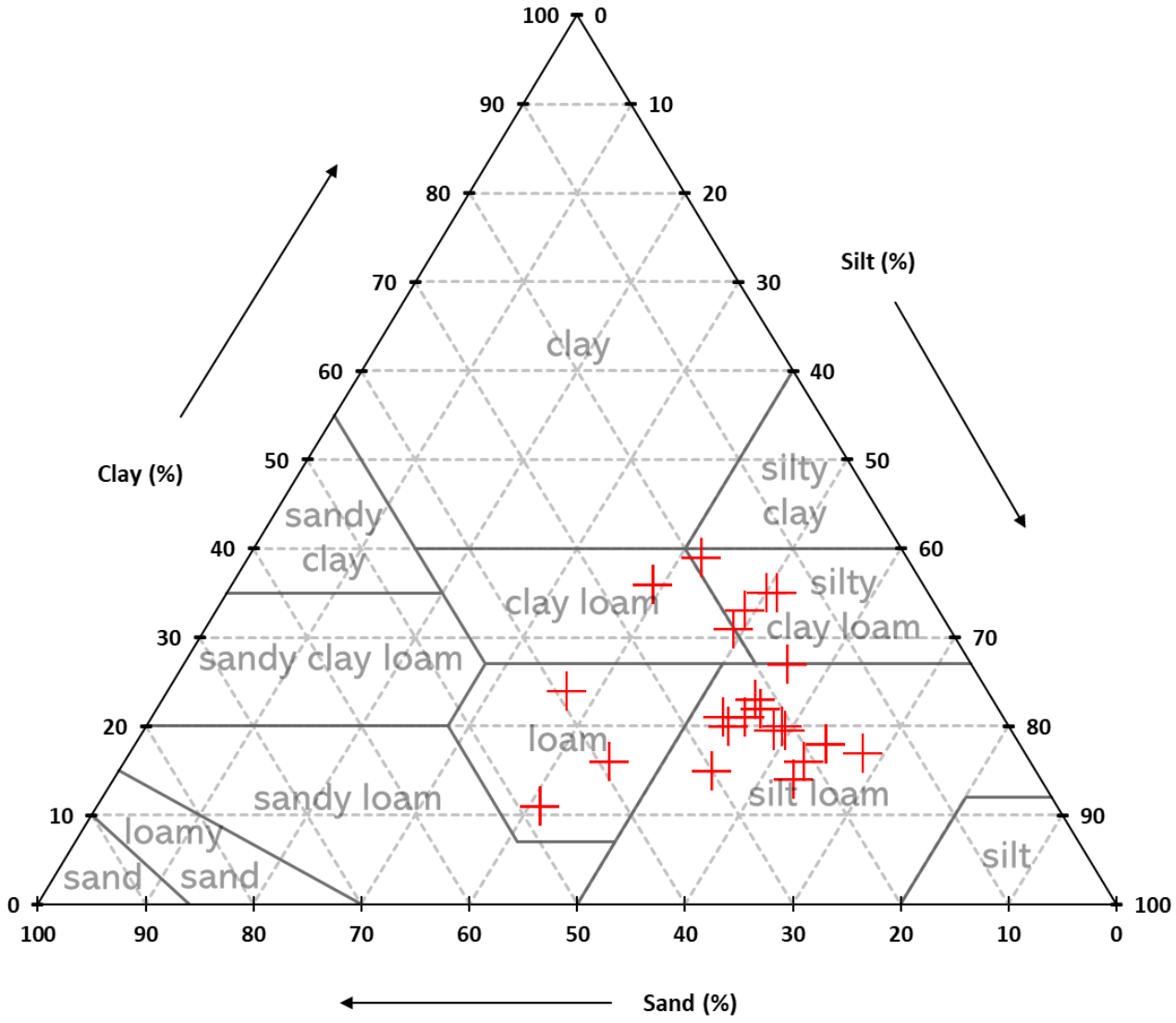


Figure 3.4. Soil texture triangle with all soil types (red cross) used in this study.

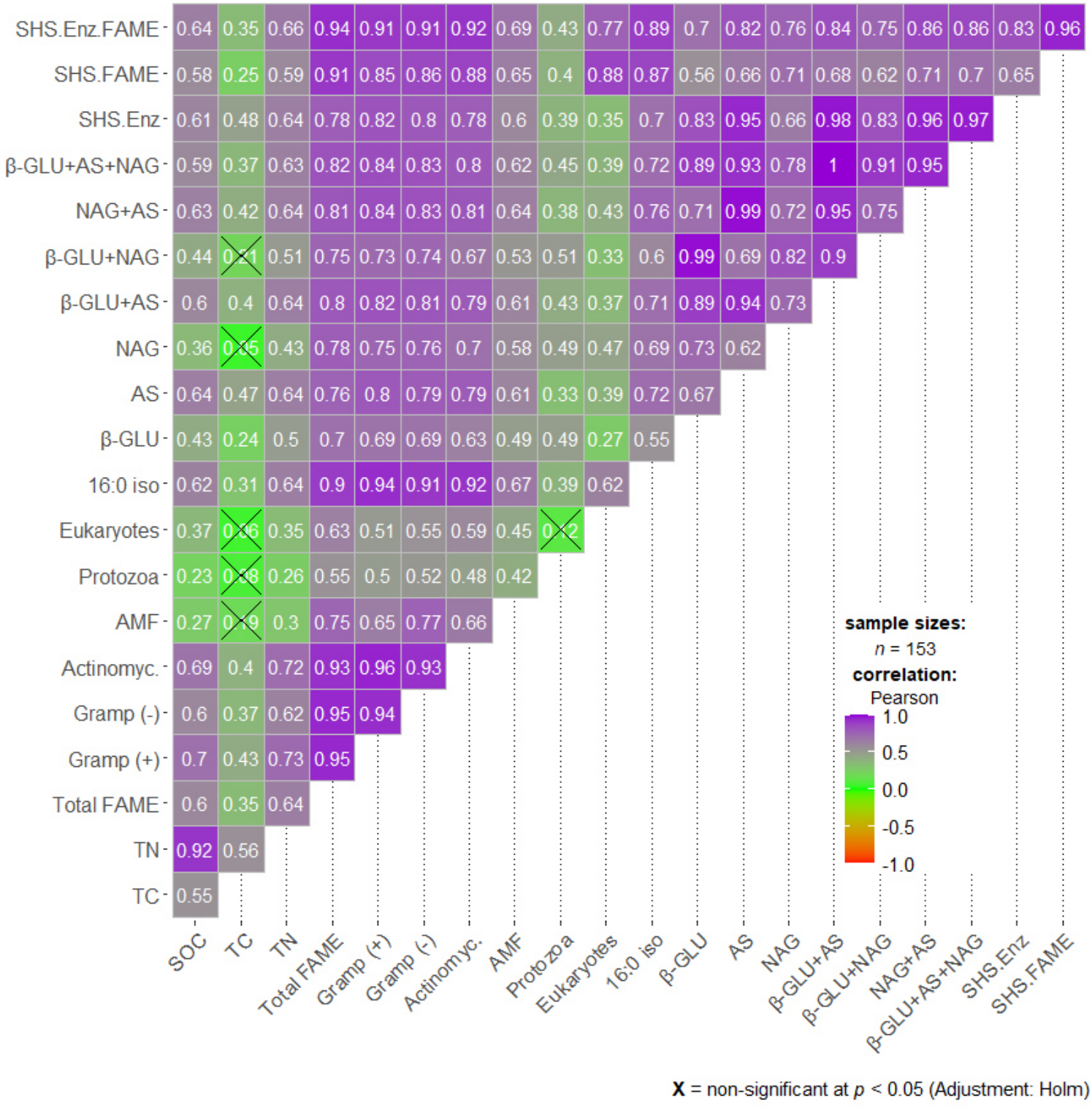


Figure 3.5. Correlation Matrix for all relevant variables in Chapter 3.4 (n=153)

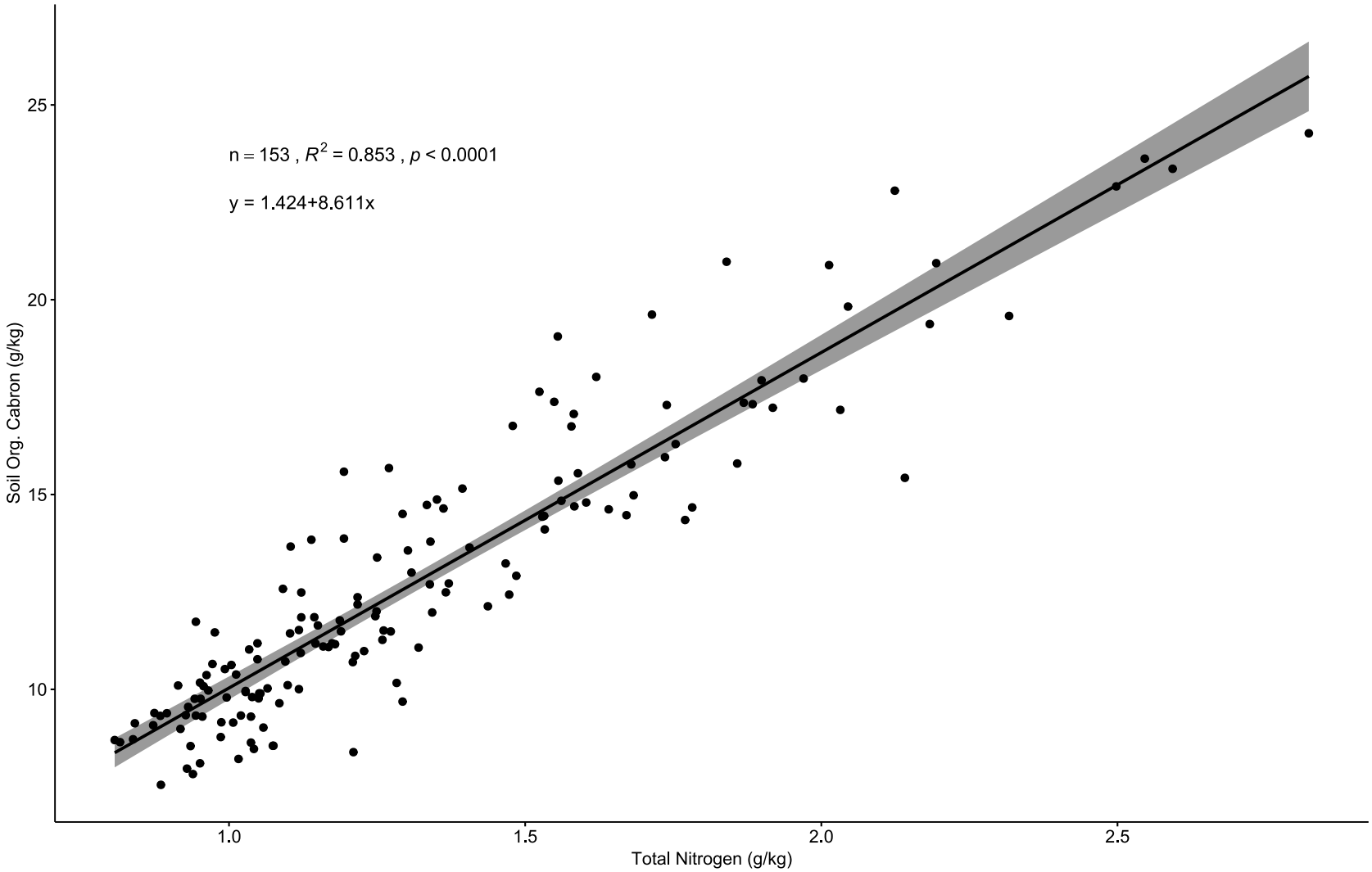


Figure 3.6. Scatter plot for Soil org. Carbon and Total Nitrogen including the linear correlation between them (n=153).

3.7 Supplementary Information

Supp. Table 3.1 List of all agricultural on-farm sites in the study. Information related to organic and conventional management practices are provided. Additionally, information related to location, seasonal soybean growing time, seasonal crop that was planted, soil texture, and soil series is provided.

ID	Farming Type	Site ID	Field ID	Elevation	State	County	Growing Period (days)			Planted Crop			Soil Texture			Soil Series
							year 1	year 2	year 3	year 1	year 2	year 3	Clay (%)	Silt (%)	Sand (%)	
1	Org	VY	F1	HE	OH	Madison	123	-	-	Soy	Wheat	Corn	18	64	18	Crosby-Lewisburg silt loams
2	Org	VY		LE	OH	Madison	123	-	-	Soy	Wheat	Corn	18	64	18	
3	Org	VY	F2	HE	OH	Madison	-	137	-	Corn	Soy	Wheat	18	64	18	Crosby-Lewisburg silt loams
4	Org	VY		LE	OH	Madison	-	137	-	Corn	Soy	Wheat	18	64	18	
5	Org	LM	F1	HE	OH	Madison	122	-	-	Soy	Wheat	Corn	18	64	18	Crosby-Lewisburg silt loams
6	Org	LM		LE	OH	Madison	122	-	-	Soy	Wheat	Corn	18	64	18	
7	Org	LM	F2	HE	OH	Madison	-	128	-	Corn	Soy	Wheat	18	64	18	Crosby-Lewisburg silt loams
8	Org	LM		LE	OH	Madison	-	128	-	Corn	Soy	Wheat	18	64	18	
9	Org	JK	F1	HE	OH	Madison	120	-	-	Soy	Wheat	Corn	22	58.5	19.5	Crosby-Lewisburg silt loams
10	Org	JK		LE	OH	Madison	120	-	-	Soy	Wheat	Corn	18	64	18	
11	Org	JK	F2	HE	OH	Madison	-	-	132	Wheat	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
12	Org	JK		LE	OH	Madison	-	125	-	Corn	Soy	Wheat	18	64	18	
13	Org	DB	F1	HE	OH	Hancock	160	-	-	Soy	Corn	Wheat	15	50	35	Pewamo silty clay loam
14	Org	DB		LE	OH	Hancock	160	-	-	Soy	Corn	Wheat	22	56	22	Blount silt loam
15	Org	DB	F2	HE	OH	Hancock	-	132	-	Corn	Soy	Wheat	39	37	24	Blount-Houcktown complex
16	Org	DB		LE	OH	Hancock	-	132	-	Corn	Soy	Wheat	22	56	22	Blount silt loam

continues

17	Org	RA	F1	HE	OH	Clinton	-	131	-	Oats / Forage	Soy	Hay / Alfalfa	17	56	27	Treaty silty clay loam
18	Org	RA		LE	OH	Clinton	-	131	-	Oats / Forage	Soy	Hay / Alfalfa	17	56	27	
19	Org	RA	F2	HE	OH	Clinton	-	-	-	Soy	Oats / Forage	Forage	17	56	27	Treaty silty clay loam
20	Org	RA		LE	OH	Clinton	-	-	-	Soy	Oats / Forage	Forage	17	56	27	
21	Conv	BH	F1	HE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	Hoytville clay loam
22	Conv	BH		LE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	
23	Conv	BH	F2	HE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	Hoytville clay loam
24	Conv	BH		LE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	
25	Conv	PD	F1	HE	OH	Fulton	-	156	-	Rye	Soy	Corn	14	51	35	Lenawee silty clay loam
26	Conv	PD		LE	OH	Fulton	141	-	151	Soy	Corn	Soy	14	51	35	
27	Conv	PD	F2	HE	OH	Fulton	134	149	-	Soy	Soy	Corn	18	49	33	Fulton silty clay loam
28	Conv	PD		LE	OH	Fulton	134	149	-	Soy	Soy	Corn	18	49	33	
29	Conv	PD	F3	HE	OH	Fulton	118	-	-	Soy	Bare	Corn	25	39	36	Hoytville clay loam
30	Conv	PD		LE	OH	Fulton	118	133	-	Soy	Soy	Corn	25	39	36	
31	Conv	PD	F4	HE	OH	Fulton	132	136	-	Soy	Soy	Corn	25	39	36	Hoytville clay loam
32	Conv	PD		LE	OH	Fulton	132	136	-	Soy	Soy	Corn	25	39	36	
33	Conv	PD	F5	HE	OH	Fulton	135	-	129	Soy	Corn	Soy	22	55	23	Del Rey silt loam
34	Conv	PD		LE	OH	Fulton	135	-	129	Soy	Corn	Soy	22	55	23	
35	Conv	PD	F6	HE	OH	Fulton	121	132	-	Soy	Soy	Corn	25	39	36	Hoytville clay loam
36	Conv	PD		LE	OH	Fulton	121	-	136	Soy	Corn	Soy	25	39	36	

continues

37	Conv	PD	F7	HE	OH	Fulton	142	-	130	Soy	Corn	Soy	22	55	23	Del Rey silt loam
38	Conv	PD		LE	OH	Fulton	142	-	130	Soy	Corn	Soy	22	55	23	
39	Conv	TL	F1	HE	OH	Morrow	134	138	-	Soy	Soy	Corn	21	63	16	Centerburg silt loam
40	Conv	TL		LE	OH	Morrow	134	138	-	Soy	Soy	Corn	21	59	20	
41	Conv	TL	F2	HE	OH	Morrow	134	-	139	Soy	Corn	Soy	21	63	16	Centerburg silt loam
42	Conv	TL		LE	OH	Morrow	134	-	139	Soy	Corn	Soy	21	63	16	
43	Conv	JM	F1	HE	OH	Hancock	-	-	152	Bare	Corn	Soy	22	56	22	Blount silt loam, ground moraine
44	Conv	JM		LE	OH	Hancock	-	-	152	Bare	Corn	Soy	22	56	22	
45	Conv	JM	F2	HE	OH	Hancock	-	140	-	Bare	Soy	Corn	22	56	22	Blount silt loam, ground moraine
46	Conv	JM		LE	OH	Hancock	-	140	-	Bare	Soy	Corn	22	56	22	
47	Conv	BG	F1	HE	OH	Hancock	-	168	-	Bare	Soy	Corn	19	42	39	Hoytville clay loam
48	Conv	BG		LE	OH	Hancock	-	168	-	Bare	Soy	Corn	19	42	39	
49	Conv	BG	F2	HE	OH	Hancock	-	-	154	Bare	Corn	Soy	19	42	39	Hoytville clay loam
50	Conv	BG		LE	OH	Hancock	-	-	154	Bare	Corn	Soy	19	42	39	
51	Conv	RB	F1	HE	OH	Darke	126	163	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
52	Conv	RB		LE	OH	Darke	126	163	-	Soy	Soy	Corn	18	64	18	
53	Conv	AO	F1	HE	OH	Darke	120	123	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
54	Conv	AO		LE	OH	Darke	120	123	-	Soy	Soy	Corn	18	64	18	
55	Conv	AO	F2	HE	OH	Darke	114	-	145	Soy	Corn	Soy	18	64	18	Crosby silt loam
56	Conv	AO		LE	OH	Darke	114	-	145	Soy	Corn	Soy	18	64	18	

continues

57	Conv	BM	F1	HE	OH	Madison	156	-	151	Soy	Corn	Soy	24	55	21	Miamian-Lewisburg silt loams
58	Conv	BM		LE	OH	Madison	156	-	151	Soy	Corn	Soy	24	55	21	
59	Conv	BM	F2	HE	OH	Madison	-	164	-	Corn	Soy	Corn	24	55	21	Miamian-Lewisburg silt loams
60	Conv	BM		LE	OH	Madison	-	164	-	Corn	Soy	Corn	24	55	21	
61	Conv	CO	F1	HE	OH	Pickaway	123	135	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
62	Conv	CO		LE	OH	Pickaway	123	135	-	Soy	Soy	Corn	18	64	18	
63	Conv	CO	F2	HE	OH	Pickaway	-	126	145	Corn	Soy	Soy	18	64	18	Crosby silt loam
64	Conv	CO		LE	OH	Pickaway	-	126	145	Corn	Soy	Soy	18	64	18	
65	Conv	CHA	F1	HE	OH	Pickaway	-	134	-	Corn	Soy	Corn	26	53	21	Miamian-Lewisburg silt loams
66	Conv	CHA		LE	OH	Pickaway	-	134	-	Corn	Soy	Corn	26	53	21	
67	Conv	CHA	F2	HE	OH	Pickaway	-	134	-	Corn	Soy	Corn	21	59.5	19.5	Crosby silt loam
68	Conv	CHA		LE	OH	Pickaway	-	134	-	Corn	Soy	Corn	18	64	18	
69	Conv	LM	F1	HE	OH	Tuscarawas	173	-	140	Soy	Corn	Soy	39	45	16	Wheeling loam
70	Conv	LM		LE	OH	Tuscarawas	173	-	140	Soy	Corn	Soy	39	45	16	
71	Conv	LM	F2	HE	OH	Tuscarawas	-	131	-	Corn	Soy	Corn	39	45	16	Wheeling loam
72	Conv	LM		LE	OH	Tuscarawas	-	131	-	Corn	Soy	Corn	39	45	16	
73	Conv	JH	F1	HE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	Crosby-Lewisburg silt loams
74	Conv	JH		LE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	
75	Conv	JH	F1 - Tra	HE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	Crosby-Lewisburg silt loams
76	Conv	JH		LE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	

continues

77	Conv	JH		HE1	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	
78	Conv	JH	F2	HE2	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
79	Conv	JH		LE	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	
80	Conv	BP	F1	HE	OH	Madison	-	-	154	Soy	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
81	Conv	BP		LE	OH	Madison	-	-	154	Soy	Corn	Soy	18	64	18	
						Mean:	133	140	143				21.3	55.2	23.5	

Suppl. Table 3.2. List of FAME biomarkers organized by taxonomic microbial group designation and fatty acid structure. Biomarkers that were used in this study are highlighted. Furthermore, information regarding their number of observations and detection is provided.

Fatty Acid	Taxonomic Group		Biomarker	References	Detected	Nr. of obs.
Saturated	-		**14:0**	-	Yes	*n=153* (100 %)
	Gram (+); Gram (-); Fungi; Eukaryotes		**16:0 / 18:0**	Zelles, 1997; Kerger, 1986	Yes	*n=153* (100 %)
	Generally considered Bacteria		**15:0 / 17:0**	Federle, 1986; Tunlid et al., 1989; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
Mid-Chain Branched	Gram Positive (+) Bacteria	Actinomycetes (Actinobacteria)	**10Me16:0 / 10Me17:0 / 10Me18:0 / 10Me19:1w7c**	Fisher et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Forstegard et al., 1993, Veum et al. 2021	Yes	*n=153* (100 %)
Terminally Branched		Gram (+)	*i14:0 **	Zelles, 1999	Yes	*n=132* (86.3 %)
			a15:0 / i15:0 / i16:0 / a17:0 / i17:0	O'Leary and Wilkinson, 1988	Yes	*n=153 (100 %)*
Mid-Chain Branched / Monounsaturated	Gram Negative (-) Bacteria	*Sulfate red. Bacteria **	10Me16:0 / cy17:0 * <u>*anoxic and anaerobic conditions*</u>	Dowling et al., 1985, 1988; Parkes et al. 1993	-	-
Hydroxy-substituted		Gram (-)	2OH 12:0 / 3OH 12:0 / 2OH 14:0 / * 3OH 14:0 / 2OH 16:0 / 2OH 18:0	Parker et al., 1982	No	*n=0* (0 %)
			17:1 w8c	Zelles, 1999	Yes	*n=153 (100 %)*
Monounsaturated			**16:1ω7c / 18:1ω7c**	Wilkinson, 1988; Tunlid et al., 1989;	Yes	*n=153 (100 %)*
			cy17:0 / cy19:0	Wilkinson, 1988; Kerger, 1986	Yes	*n=153 (100 %)*
	Methanogens * (anaerobic conditions)		Type I: 16:1ω5t / 16:1ω7c / 16:1ω8c / 3OH 16:0 *	Nichols et al. 1985; Bowman et al. 1991, 1993	-	-
			Type II: 18:1ω8c / 18:1ω7c *		-	-

continues

Monounsaturated	Fungi	Fungi; *Plants* *	20:1ω9c *	Madan et al., 2002	Yes	*n=92* (60.1 %)
		Arbuscular Mycorrhizal Fungi (AMF); *Plants*	**16:1ω5c**	Olsson et al., 1995; Olsson, 1999; Madan et al., 2002	Yes	*n=153 (100 %)*
Polyunsaturated		Saprophytic Fungi; P*lants*	**18:1ω9c**	Vestal and White, 1989; Wallis et al. 2021	Yes	*n=153 (100 %)*
			18:2ω6c	Federle, 1986; Zelles, 1997; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
			18:3ω6c *	Federle, 1986; Klug, 1996	Yes	*n=140* (91.5 %)
		Fungi; *Plants* *	18:3ω3c *	Zelles, 1997	No	*n=0* (0 %)
		Fungi; *Plants* *	20:5ω3c *	Nordby et al., 1981; Olsson et al., 1995	Yes	*n=109* (71.2 %)
	Protozoa		**20:3ω6c**	Nordby et al., 1981; Guckert et al., 1985	Yes	*n=13* (8.5 %)
			20:4ω6c		Yes	*n=153* (100 %)
Saturated (Long Chain)	Eukaryotes		**21:0 / 22:0 / 24:0**	Zelles, 1999	Yes	*n=153* (100 %)
			23:0		Yes	*n=98* (64.1 %)

* Biomarkers were not classified in this study as such due to different environmental conditions or number of observations

Suppl. Table 3.3. Enzyme Activity Assay protocols for the individual enzymes (β-GLU, AS, NAG).

Enzyme Assay description and ecological role	Incubation step (37 °C for 1h)		After Incubation	
	Buffer	Substrate [a)]	$CaCl_2$ (0.5 M)	Stop solution
β-glucosidase (C cycling)	MUB pH 6.0; 4 mL	p-Nitrophenyl-β-D-glucopyranoside (0.05 M); 1 mL Sigma N7006	1 mL	THAM pH 12 (0.1 M); 4 mL
Arylsulfatase (S cycling)	0.5 M Acetate Buffer pH 5.8; 4 mL	p-Nitrophenyl sulfate (0.05 M); 1 mL; Sigma N3877	1 mL	THAM pH 12 (0.1 M); 4 mL
β-glucosaminidase (C and N cycling)	0.1 M Acetate Buffer pH 5.5; 4 mL	p-Nitrophenyl-N-acetyl-β-D-glucosaminide (0.01 M); 1 mL; Sigma N9376	1 mL	THAM pH 12 (0.1 M); 4 mL

a) Substrates were prepared by using the corresponding incubation buffer (Tabatabai, 1994).

Suppl. Table 3.4. Variables used in the multivariate model analysis runs. Variables are grouped into predefined classifications. These classifications are separated into groups like environmental factors, agricultural land management factors, and groups related to soil specific properties.

Basic Factors

		Classification	*Variable (x_i)*	*Unit*
Basic Factors	**Type**	Organic or Conventional Farming	isOrganic	1 or 0
	Texture	Sand (0.05 - 0.002 mm)	Sand	%
		Silt (0.05 - 2 mm)	Silt	%
		Clay (<0.002 mm)	Clay	%

Tier 1 Soil Health Indicators

	Classification	*Variable (x_i)*	*Unit*
Tier 1 Indicator (Core)	Soil Organic Carbon	SOC	N %
	Total Nitrogen	TN	C %

Soil Chemical Factors

		Classification		*Variable (x_i)*	*Unit*
Tier 1 Indicator (Chemical)	**pH**	pH		pH	-
	pH Range	6.0 – 7.0		pH ideal	1 or 0
		5.8 - 6.0	7.0 -7.4	pH very good	1 or 0
		5.4 - 5.8	7.4 - 7.8	pH good	1 or 0
		5.0 - 5.4	7.8 - 8.2	pH ok	1 or 0
		< 5.0	> 8.2	pH bad	1 or 0

Environmental Factors

		Classification	*Variable (x_i)*	*Unit*
Environmental Factors	**Growing Cycle**	Precipitation during growth season	Precip	mm
		Growing Time (planting to harvest)	Time	days
	Soil Order	Mollisol or Alfisol	Mollisol	1 or 0

Land Management Factors

		Classification	*Variable (x_i)*	*Unit*
Agricultural Land Management Factors	**Cover Crop usage**	CC usage once in 3 years	CC13y	1 or 0
		CC usage twice in 3 years	CC23y	1 or 0
		CC usage every year	CC33y	1 or 0
		No CC was planted	No CC	1 or 0
	Pesticide usage	No Herbicide (primarily org. Farming)	HerbNo	1 or 0
		Only Glyphosate	HerbGlyp	1 or 0
		Glyph. Mix	HerbMix	1 or 0
		Glyph. Mix + Dicamba	HerbDicamba	1 or 0
		Fungicide (Yes or No)	Fungicide	1 or 0
	Manure usage	Manure usage once in 3 years	Manure13y	1 or 0
		Manure usage twice in 3 years	Manure23y	1 or 0
		Manure usage every year	Manure33y	1 or 0
		No Manure was used	No-Manure	1 or 0
		Chicken Manure usage	usesChM	1 or 0
		Cattle Manure usage	usesCM	1 or 0

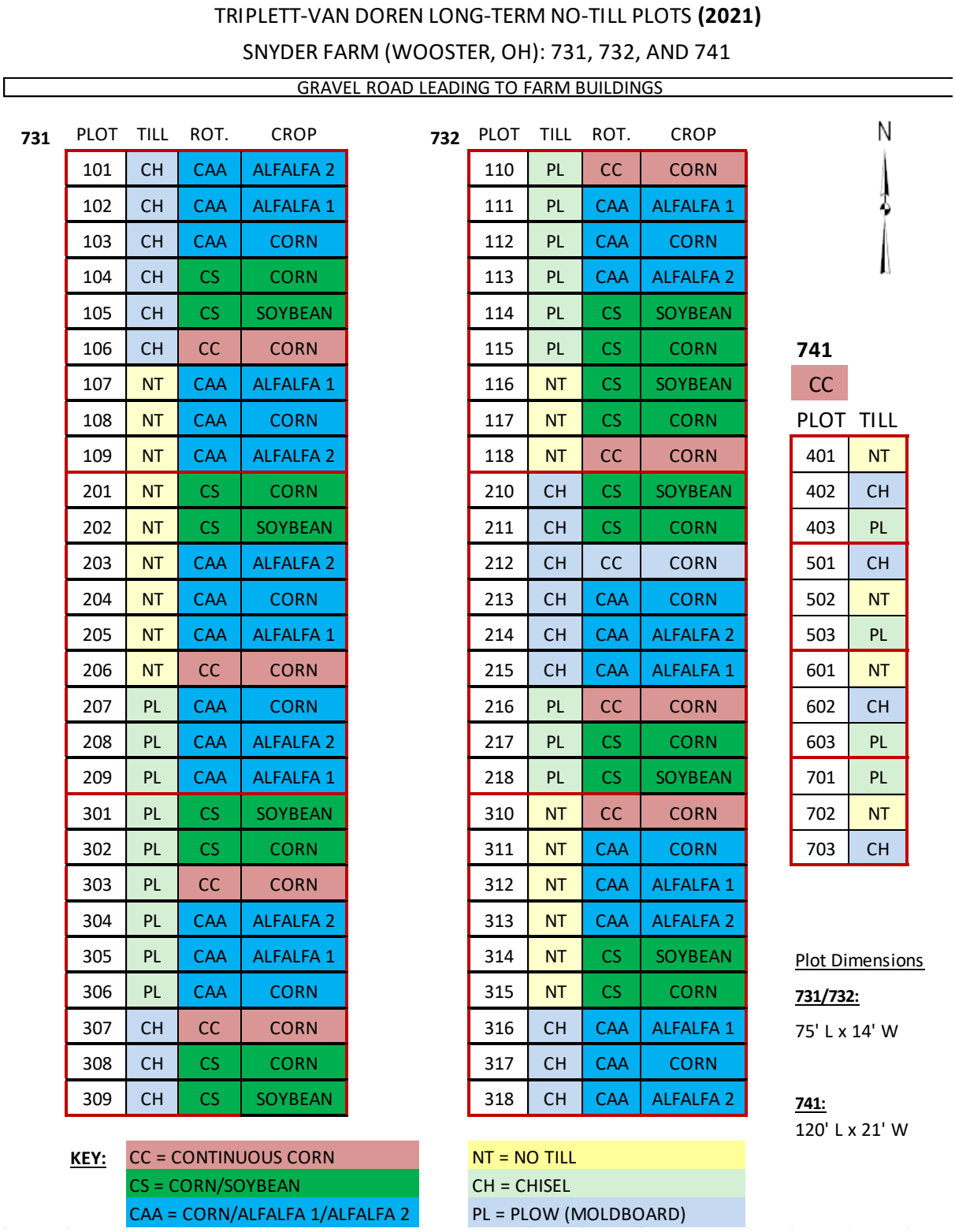
Land Management Factors

		Classification	*Variable (x_i)*	*Unit*
Agricultural Land Management Factors	**Planting date**	April date	Early	1 or 0
		May date	Common	1 or 0
		June date	Late	1 or 0
	Crop Rotation	Corn-Soybean	RotCS	1 or 0
		Corn-Soybean-Wheat	RotCSW	1 or 0
	Tillage Practice	No-Tillage	TillNT	1 or 0
		Chisel Tillage	TillChis	1 or 0
		Chisel + Field Cultivator	TillChFC	1 or 0
		Disk Tillage	TillDisk	1 or 0
		Moldboard Plow	TillMBplow	1 or 0
	Surface Residue Coverage	Residue Coverage	Residue	%
		> 60% Coverage	ResHigh	1 or 0
		30 - 60 % Coverage	ResMid_Plus	1 or 0
		15 - 30 % Coverage	ResMid	1 or 0
		6 - 15 % Coverage	ResLow Plus	1 or 0
		< 6% Coverage	ResLow	1 or 0

continues

Soil Biochemical / Biological Factors

		Classification	*Variable (x_i)*	*Unit*
Enzyme	**Enzyme Activity**	β-Glucosidase (GLU)	GLU	mg PNP / kg- dry soil * h
		N-Acetyl Glutamate synthase (NAG)	NAG	mg PNP / kg- dry soil * h
		Arylsulfatase (AS)	AS	mg PNP / kg- dry soil * h
		GLU + AS	GLUAS	mg PNP / kg- dry soil * h
		GLU + NAG	GLUNAG	mg PNP / kg- dry soil * h
		NAG + AS	NAGAS	mg PNP / kg- dry soil * h
		GLU + NAG + AS	GLUNAGAS	mg PNP / kg- dry soil * h
	Enzyme Ratios	Enzyme Ratio	GLU / AS	-
		Enzyme Ratio	GLU / NAG	-
		Enzyme Ratio	NAG / AS	-
		Enzyme Ratio	(GLU + AS) / NAG	-
		Enzyme Ratio	(GLU + NAG) / AS	-
		Enzyme Ratio	(AS + NAG) / GLU	-
EL-FAME	**FAME Biomarkers**	*Only measurable and reappearing FAME biomarkers were used in the Model development* ***(32 biomarker variables)***		nmol / g- dry soil
		Summation	Actinobacteria	nmol / g- dry soil
		Summation	Eukaryotes	nmol / g- dry soil
		Summation	Fungi	nmol / g- dry soil
		Summation	Gram- bacteria	nmol / g- dry soil
		Summation	Gram+ bacteria	nmol / g- dry soil
		Summation	Protozoa	nmol / g- dry soil
		Summation	Total Biomarker	nmol / g- dry soil
		Summation	Total Fungi	nmol / g- dry soil
	FAME Ratio	FAME Ratio	Bacteria / Total FAME	-
		FAME Ratio	Fungi / Total FAME	-
		FAME Ratio	Fungi / Bacteria	-
		FAME Ratio	AMF / Bacteria	-
		FAME Ratio	SAT / MONOSAT	-
		FAME Ratio	Gram+ / Gram-	-
		FAME Ratio	cy17:0 / Precursor (16:1ω7c)	-
		FAME Ratio	cy19:0 / Precursor (18:1ω7c)	-



Suppl. Figure 1. 2021 Plot map for the Triplett-Van Doren Long-Term Field Site in Wooster, Ohio.

LONG-TERM TILLAGE PLOTS (2021)
NORTHWEST RESEARCH STATION (Hoytville, OH): FieldTA - 3

N

PLOT	TILL	ROT.	CROP
101	CH	CAA	CORN
102	CH	CAA	ALFALFA 2
103	CH	CAA	ALFALFA 1
104	NT	CS	SOYBEAN
105	NT	CS	CORN
106	PL	CAA	CORN
107	PL	CAA	ALFALFA 2
108	PL	CAA	ALFALFA 1
109	PL	CC	CORN
110	CH	CS	SOYBEAN
111	CH	CS	CORN
112	NT	CC	CORN
113	NT	CAA	CORN
114	NT	CAA	ALFALFA 2
115	NT	CAA	ALFALFA 1
116	PL	CS	SOYBEAN
117	PL	CS	CORN
118	CH	CC	CORN

PLOT	TILL	ROT.	CROP
201	PL	CS	CORN
202	PL	CS	SOYBEAN
203	CH	CC	CORN
204	NT	CAA	ALFALFA 2
205	NT	CAA	CORN
206	NT	CAA	ALFALFA 1
207	PL	CC	CORN
208	CH	CS	CORN
209	CH	CS	SOYBEAN
210	CH	CAA	ALFALFA 2
211	CH	CAA	CORN
212	CH	CAA	ALFALFA 1
213	NT	CC	CORN
214	NT	CS	CORN
215	NT	CS	SOYBEAN
216	PL	CAA	ALFALFA 2
217	PL	CAA	CORN
218	PL	CAA	ALFALFA 1

PLOT	TILL	ROT.	CROP
301	NT	CC	CORN
302	PL	CC	CORN
303	NT	CS	SOYBEAN
304	NT	CS	CORN
305	CH	CAA	ALFALFA 1
306	CH	CAA	CORN
307	CH	CAA	ALFALFA 2
308	PL	CS	SOYBEAN
309	PL	CS	CORN
310	CH	CC	CORN
311	PL	CAA	ALFALFA 1
312	PL	CAA	CORN
313	PL	CAA	ALFALFA 2
314	CH	CS	SOYBEAN
315	CH	CS	CORN
316	NT	CAA	ALFALFA 1
317	NT	CAA	CORN
318	NT	CAA	ALFALFA 2

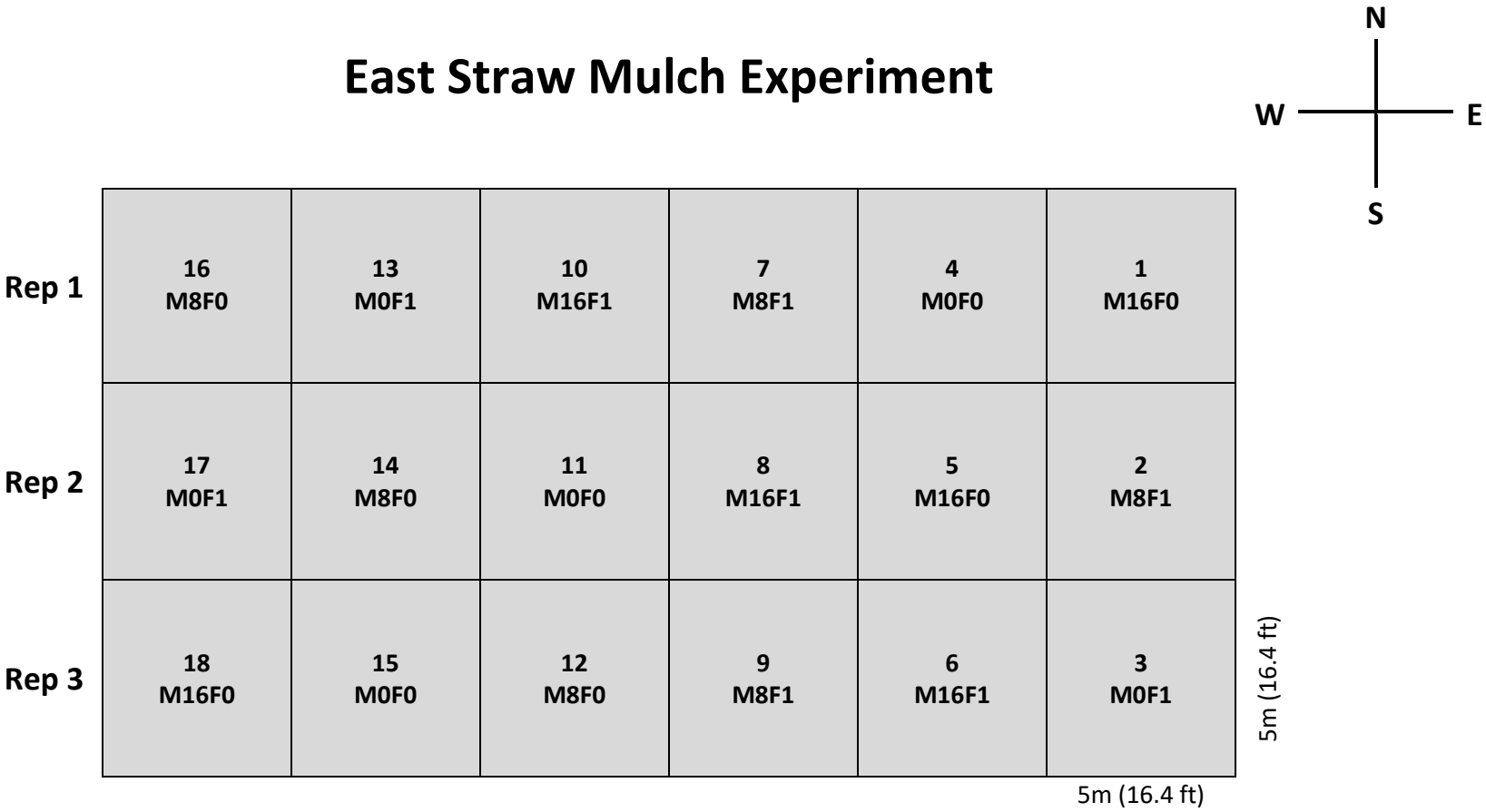
KEY:

CC = CONTINUOUS CORN
CS = CORN/SOYBEAN
CAA = CORN/ALFALFA 1/ALFALFA 2

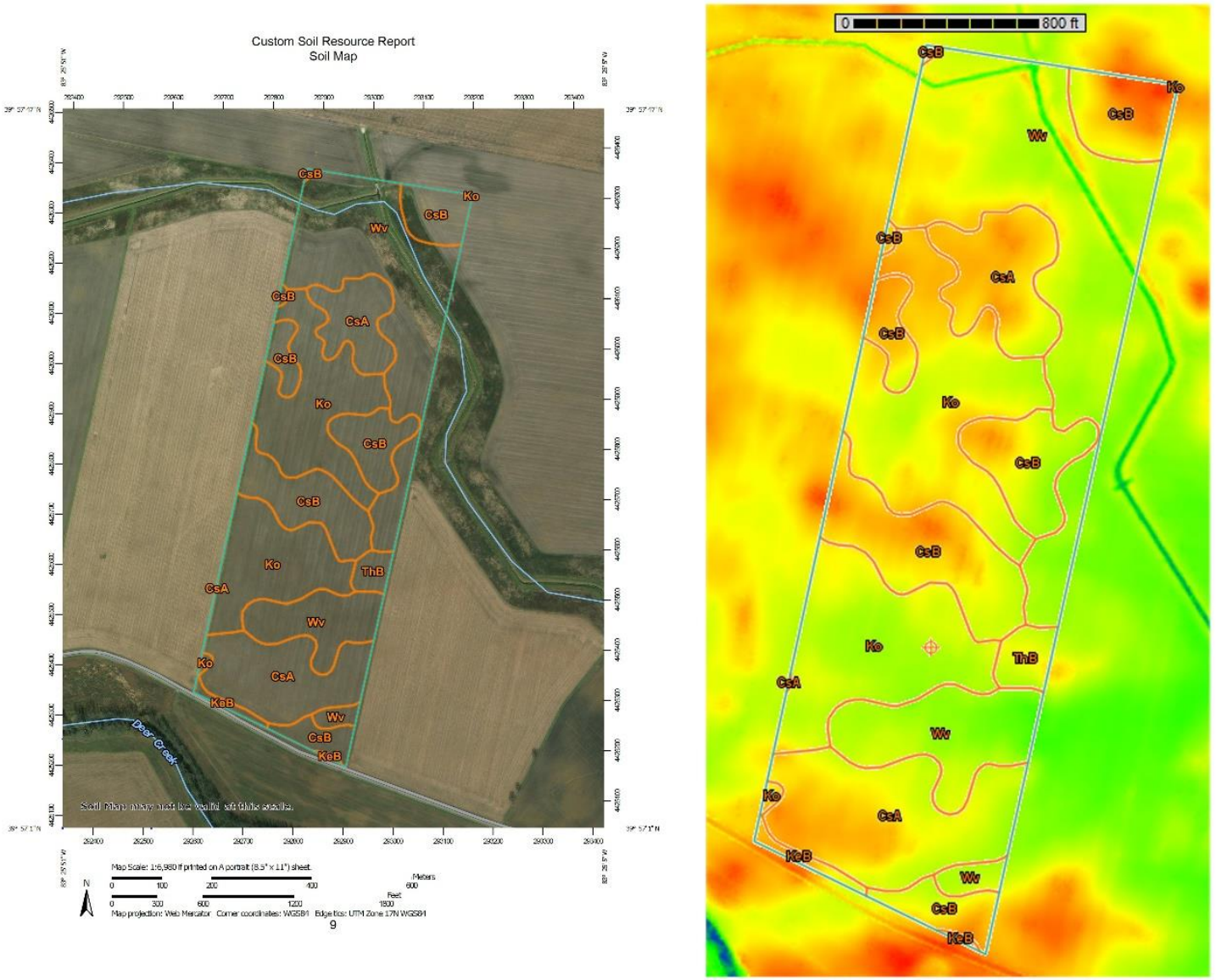
NT = NO TILL
CH = CHISEL
PL = PLOW (MOLDBOARD)

PLOTS: 20' W x 90' L

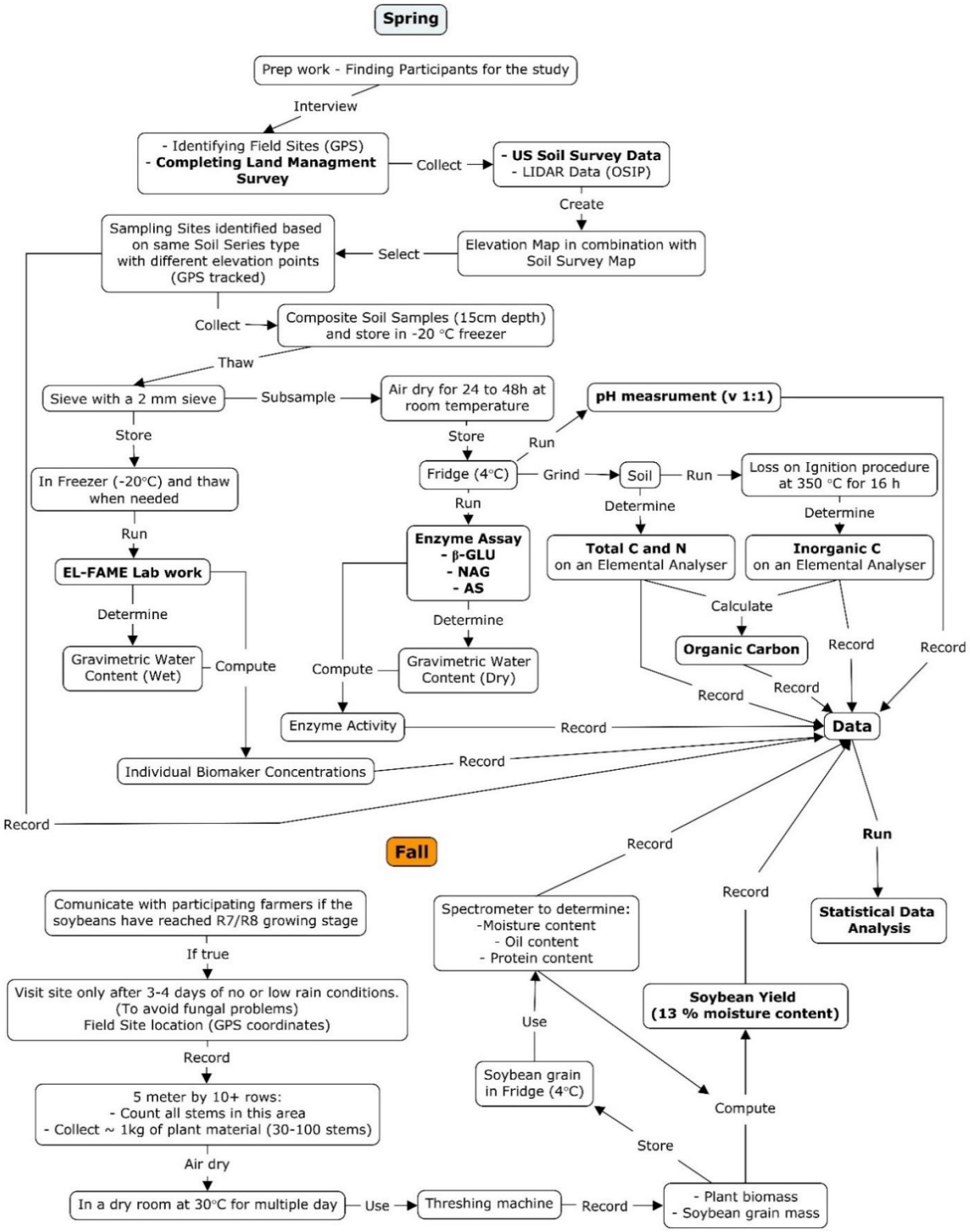
Suppl. Figure 2. 2021 Plot map for the Northwest Research Station in Hoytville, Ohio.



Suppl. Figure 3. 2021 Plot map for the East Straw Mulch Experiment in Columbus, Ohio.



Suppl. Figure 4. Example of a soil map overlayed on top of an elevation heat map created with LIDAR data



Suppl. Figure 3.5. Flow Chart of the experimental study design representing soil related measurements in spring and soybean yield related measurement steps in fall.

Chapter 4: The Ability of Soil Properties and the Novel Biochemical Soil Health Scores to Differentiate between Agricultural Management Systems

4.1 Abstract

Soil Health (SH) indicators need to be identified that can quantitatively detect effects of agricultural systems and guide sustainable management of soils. Roper et al. (2017) showed that existing SH tests and scores had limited ability to distinguish sustainable soil management systems (e.g. no-till, organic amendments) from systems that degrade soils (e.g. moldboard plowing). Therefore, the objective was to investigate the potential of soil microbial indicators, enzyme assays, soil organic carbon (SOC), and total nitrogen (TN) to differentiate among agricultural management systems at on-farm sites in Ohio, and four long-term field sites. Practices varied in crop rotation, tillage, cover cropping, and manure amendments. Additionally, over two years, non-agricultural soil samples were taken once/year at restored prairies and unmanaged, virgin soil sites. Soil samples (0-15 cm) were measured once/year over three growing seasons at the agricultural sites. A sensitivity analysis of the 521 variables identified 30 SH indicators. A secondary correlation analysis between SOC reduced the number of SH indicators to eight. Two of them were ester linked fatty acid methyl ester (EL-FAME) biomarkers (16:0 iso, 18:0 10-ME). The remaining ones were the EL-FAME marker for the soil bacterial community (sum of 11 biomarkers), three SH score variables based on EL-FAME variables, and two SH scores based on EL-FAME and enzyme activity variables. An analysis of the 8 SH indicators showed that non-agricultural systems had significantly greater EL-FAME concentrations and SH scores than soils from agricultural sites with corn-soy (CS) and corn-soy-wheat

(CSW) rotations. Significantly greater EL-FAME concentrations and SH scores were determined for cover cropping, and manure amendment. Intensive (PT), chisel (minimal) (MT), and no-tillage (NT) practices could not be differentiated by the 8 SH indicators even at low, medium, or high SOC concentrations. A deeper analysis when crop rotation, cover cropping and manure amended systems were combined with tillage showed that PT without cover cropping and manure in a CS rotation resulted in the lowest SH scores. NT with cover cropping and manure in a CSW rotation resulted in the highest SH scores. The results for the Michigan LTES had SH scores that, on average, were comparatively lower than the Ohio sites. This observation provides evidence that biological soil properties vary as function of soil type and the local climate environment. This confounds the potential of SH indicators to have absolute SH indicators that are universally interpretable across different regions and soil types. This in part can account for the inconsistent results found in previous studies on SH indicators.

4.2 Introduction

The health of a soil has a significant influence on the productivity and ecological health of a given ecosystem. Soil is a nonrenewable natural resource that must be protected to ensure food security for existing and future generations. However, quantifying soil health (SH) and identifying SH indicators has not been successfully accomplished. Previous studies have shown that several widely promoted SH tests or scores had limited ability to detect agricultural management effects, and correlated fair to poorly with crop yields (Roper et al., 2017, for three SH tests and

crop yield; Chu et al., 2019, for the Haney SH test; Chahal and Eerd, 2018, for the Haney SH test and crop yield; Faé et al., 2020, CASH SH test and soybean yield). The Soil Health Institute conducted a massive study in 2019 that involved 124 long-term experimental research sites (LTES) around North America (Norris, et al 2020). After the study was evaluated, the Soil Health Institute has recommended three SH indicators, which are soil organic carbon (SOC), carbon mineralization potential, and aggregate stability (Liptzin et al., 2022; Liptzin et al., 2023). In Chapter 3 a multivariate soybean yield prediction model, and two SH indicators that could detect soil management effects at four LTES were successfully identified. However, identifying SH indicators to soil management effects at on-farm fields are relatively few.

A meta-analysis study indicated that biological soil properties like fungal biomass, soil microbial biomass, and enzyme activities are temporally sensitive (1 to 3 years) to land management at a short time scale (Stewart et al., 2018). Studies have also shown that enzyme activities and the microbial indicators are affected by cover cropping (Bandick and Dick, 1999; Schutter and Dick, 2002), organic amendments (Carlson et al., 2015; Dick et al., 1988), climate effects (Acosta-Martínez et a., 2014 a, b), and tillage (Lorenz et al. 2020; Montero et al., 2004; Acosta-Martínez et al., 2003; Frey et al., 1999; Deng and Tabatabai, 1997; Dick, 1986a; Dick, 1986b; Dick, 1984; Doran, 1980).

Mechanistically enzymes are important in driving the carbon, nitrogen, and sulfur nutrient cycles in soils. The β-glucosidase (GLU) enzyme hydrolyzes cellulose to produce glucose, a critical energy source for the microbial community in the soil

(Tabatabai, 1994). The enzyme arylsulfatase (AS) is involved in the sulfur cycle by releasing plant available SO_4^{2-}. Because fungal communities contain ester sulfate compounds AS is highly correlated with fungal biomass (Miller et al., 1998, Bandick and Dick, 1999).

These studies have mostly been conducted on long-term experimental research plots and not soils from farmer fields. On-farm research of SH indicators is confounded by each farm site having unique management practices, that additionally may vary over time, which complicates replication across farms. Furthermore, activities of specific enzymes or microbial indicators may change depending on the composition of amendments, the relative availability of nutrients, as well as other factors, such as soil type and other physiochemical characteristics (Acosta-Martínez et al., 2007 a, b; Huang et al., 2022).

In these situations, clay content becomes important because for certain enzymes a significant amount of their activities is coming from the abiontic enzyme fraction which are enzymes stabilized in the soil matrix (clays in particular) that catalytically are no longer under the control of soil microorganisms (Skujiņš, 1978). Mandal et al. (2022) reported on the role of clays and other colloids the persistence and stability of enzymes by showing that adsorption of several enzymes (including β-glucosidase) on clay increased their activities. Other studies have shown that β-glucosidase and arylsulfatase have a significant amount of abiontic enzyme activity (Klose et al., 1999, for arylsulfatase; Knight and Dick, 2004, for β-glucosidase). Knight and Dick (2004) reported that activity from the abiontic fraction of β-glucosidase was affected by the soil management system, but this was not the case for

the activity associated with viable cells. This is an important finding, in that normalizing for clay content enables an SH assessment independent of soil type; overcoming the limits of nearly or likely all, potential chemical, physical and biological measures of SH, where soil type varies more than the more subtle effects of soil/crop management. However, this needs to be tested on a much wider range of soils, environments, and land management systems, and on soils under farmer management.

Therefore, the objectives of this study were: 1) determine the sensitivity of 521 soil variables for detecting crop/soil management using a multi-level sensitivity test in which the level of subcategories increases; 2) to compare SH indicator analyses of farm fields with the unmanaged ecosystem soils of converted conservation reserve program (CRP) sites, restored prairies (P) and untouched virgin soils; and 3) to rate the soil health of various soil/crop farmer practices based on the most sensitive SH indicators identified in this research. To identify the most sensitive SH indicators, it was hypothesized that an indicator could differentiate between non-agricultural and agricultural management effects, is temporarily sensitive, and has a correlation coefficient (r) with SOC that is greater than 0.5. To identify beneficial soil management practices, it was hypothesized that a sensitive SH indicator would result in statistically significant higher enzyme activities, C and N concentrations, or SH scores due to crop rotation, cover cropping, manure amendments, or tillage practices.

4.3 Material and Methods

4.3.1 Study Sites

A total of 301 soil samples were collected. One hundred and seventy-six originated from conventionally managed on-farm sites (~58.5%), 52 from organically managed on-farm sites (~17%), 48 from long-term field sites (LTFS) (~16%), 16 from restored prairie sites (~5.5%), and the remaining 9 (~3%) were collected at two separate virgin soil sites (untouched soils). On-Farm soil sampling occurred each spring in a period of three years (2019-2021). For restored prairies and virgin soils sampling was done for two years (2020-2021), and in 2021 soil samples were collected at the four LTES. Three of the LTES are located in Ohio and one in Michigan. The restored prairies and two virgin soils sites are located in Madison county, Ohio (Suppl. Figure 7). Each sample site was GPS-tracked. Most samples were collected at on-farm sites (n=228, ~76%). These samples are connected to 18 farmers, which were located throughout eight counties in Ohio (Clinton, Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas). From the 176 original conventionally managed on-farm sites 6 were classified as conservation reservation program (CRP) sites. These sites were in the CRP for close to 25 years and were turned back into agricultural farmland sometime between 2019 to 2020.

4.3.2 Farm Study Sites

4.3.2.1 Organic Farm Sites

Fifty-two soil samples originated from certified organically managed farm field sites in Ohio. The organic farm fields were in Madison, Handcock, and Clinton county and have been under agricultural practices for 50 to more than 100 years. Fields under

organic management had been in place from to one to 20 years and range in size from 5 to 47 ha. Precipitation for each site ranged from 244 to 553 mm (Climate Fieldview, n.d). The growing period from planting to harvest ranged from 120 to 160 days. Most organically managed soils were a Crosby-Lewisburg silt loams (mesic Aeric Epiaqualfs / shallow Aquic Hapludalfs) and four were Mollisols (Soil Survey Staff, 2019). Furthermore, the only four Mollisols in the study were identified at two separate organic farm field locations. All organic farm sites used organic seeds, had a soybean-corn-wheat rotation, a 30-inch (76 cm) row spacing, and no synthetic inputs to meet certified organic standards. However, across the organically, managed fields there was variation in tillage manure applications, and cover cropping.

4.3.2.2 Conventional Farm Sites

The 13 conventionally managed fields ranged in size from 4 to 77 ha. Twenty-eight (~ 16%) had a soybean-corn-wheat rotation and 148 (~ 84%) a soy-corn rotation. All conventional farmers used synthetic fertilizers and herbicides. Herbicide management was divided in three application categories: only glyphosate (*N-(phosphonomethyl) glycine*), glyphosate with a secondary herbicide, and dicamba (*3,6-dichloro-2-methoxybenzoic acid).* All conventionally managed fields had Alfisols (Soil Survey Staff, 2019). The soil samples came from Darke, Fulton, Hancock, Madison, Morrow, Pickaway, and Tuscarawas counties. The agricultural fields ranged usage from 1 to 100 years. One field was converted from native land to farmland.

4.3.3 Long Term Field Sites

4.3.3.1 Wooster - Triplett-Van Doren Site

The LTES in Wooster, OH (40.764° N, -81.906° W) was established in 1962 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Wooster silt loam (fine-loamy, mixed, active, mesic Oxyaquic Fragiudalfs) with a 2-6 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25-30 % for sand, 55-60 % for silt and 15% for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). Deiss et al. (2021) reported a range of 5.4 to 6.8 for soil pH.

The experimental has a two-way factorial randomized complete block design with three replications with three tillage treatments, and three crop rotations (Dick and Van Doren Jr., 1985; Deiss et al., 2021). Plot size is 22.3 m by 4.3 m.

The three tillage treatments are: (1) no-tillage (NT); (2) chisel (minimum) tillage (CT); or (3) moldboard plow (MP). The minimum tillage treatment had a para plow from 1962 to 1984, after which a chisel cultivator was used. The chisel tillage loosen the soil and allows up to 30% litter retention on the soil surface. Moldboard tillage inverts soil to a depth of 20 cm and buries the litter, leaving 5 % or less on the soil surface (Dick et al., 2013).

The three crop rotation treatments on the site are: (1) continuous corn *(Zea mays L.)* (CC); (2) corn and soybean *(Glycine max L.)* (CS); and (3) corn and oat *(Avena sativa L.)* and/or alfalfa *(Medicago sativa)* or clover *(Trifolium repens L.)* (CFF). Nine soil samples were collected in 2021 from the CS rotation plots that grew soybeans (Suppl. Fig. 1.).

4.3.3.2 Hoytville - Triplett-Van Doren Site

The LTES in Hoytville, OH (41.222 ° N, -83.762° W) was established in 1963 by Glover B. Triplett and David M. Van Doren. The primary soil series is a Hoytville clay loam (fine, illitic, mesic Mollic Epiaqualfs) with a 0-1 % slope. For the first 15 cm the soil particle size distribution (texture) ranges between 25 % for sand, 39 % for silt and 36 % for clay (Dick and Van Doren Jr., 1985; Dick et al., 1986a; Soil Survey Staff, 2019). In contrast to the Wooster soil, The Hoytville soil has a poor surface and internal drainage, and it cracks when dry. In 1952 a subsurface tile drainage was installed at a depth of 1.2 - 1.4 m (Dick et al., 1986b; Deiss et al., 2021). Deiss et al. (2021) reported a range of 4.3 to 7.5 for soil pH.

It has a two-way factorial randomized complete block design with three replications, and the identical three tillage treatments, and three crop rotations as the Wooster LTES (Dick and Van Doren Jr., 1985; Deiss et al., 2021). The plot size is 30.5 m by 6.4 m. Nice soil samples were collected in 2021 from the CS rotation plots that grew soybeans (Suppl. Fig. 2.).

4.3.3.3 Columbus - Straw Mulch Experiment

The Straw Mulch Experiment (40.017° N, -83.0395° W) was established in 1996 by the Carbon Management and Sequestration Center (CMASC) at the Ohio State University. The objective of this LTES is to determine the effect of wheat straw *(Triticum aestivum L.)* mulching on soil quality, soil organic carbon (SOC) sequestration and dynamics, and greenhouse gas emissions (Blanco-Canqui and Lal, 2007). No mechanical tillage is used, and glyphosate *(N-Phosphonomethyl glycine)* is

used to control weeds. The primary soil series is a Crosby silt loam (fine, mixed, active, mesic Aeric Epiaqualfs) with a 2-6 % slope (Soil Survey Staff, 2019). For the top 15 cm the soil particle size is 22-23 % for sand, 53-56 % for silt, and 22-24 % for clay (Soil Survey Staff, 2019; Nawaz et al., 2016; Saroa and Lal, 2003). Measured soil pH at a depth of 0 to 15 cm ranged from 5.7 to 7.1.

The experimental design is a two-way factorial completely randomized block design (3 replications) with three mulch rates and two fertilizer rates. The fry mulch treatments are: (1) no mulch (control), (2) 8 Mg ha^{-1} yr^{-1}, and (3) 16 Mg ha^{-1} yr^{-1}. The fertilizer treatments are: (1) no fertilizer application (control), or (2) annual broadcast fertilizer application with a rate of 244 kg N ha^{-1} (184 kg N ha^{-1} as Urea) and 60 kg ha^{-1} of NPK). Each year, the wheat straw is applied in the spring followed by fertilizer application in the late spring to early summer. Until 2020 no crops were grown on the plots after which for two years corn and soybean were grown on them. Plot size is 5 by 5 m. Each plot on which the crop experiment took place was separated into two halves (2.5 by 5 m) with a corn-corn and soybean-soybean rotation. For this study only six soil samples were collected originating from plots with no fertilizer application and low (0 Mg/ha) and high (16 Mg/ha) mulch rates that had soybeans grown on them (Suppl. Fig. 3.).

4.3.3.4 Michigan - KBS Long-Term Ecological Research Station

The Kellogg Biological Station Long-Term Ecological Research project was established in 1987 by Michigan State University and is funded by the National Science Foundation and by the Michigan State University AgBioResearch program. Soil

samples were collected from the Main Cropping System Experiment (42.410° N, -85.373° W) which was completed in 1989. The primary soil series is a Kalamazoo loam (fine-loamy, mixed, active, mesic Typic Hapludalfs) with a 2-6 % slope. For the top 15 cm the soil particle size distribution is 32 - 50 % for sand, 34 - 39 % for silt and around 11-19 % for clay (Robertson et al., 2020; Soil Survey Staff, 2019). The soil pH in the 0 to 15 cm ranges from 5.7 to 6.5. The plot size is 87 by 105 m.

It has a factorial randomized complete block design with six replications. The tillage treatments are: (1) conventional chisel (minimum) tillage (MT-Conv); (2) conventional no-tillage (NT-Conv); (3) chisel tillage with reduced- N input (MT-Conv(-N)); and (4) biologically (organic) based system with chisel tillage (MT-Org) (Martin and Sprunger, 2022; Naasko et al. 2024). The four tillage treatments follow a corn-soybean-wheat rotation, but winter cover crops are incorporated in the reduced input (MT-Org) and biologically based systems (MT-Org) following corn and soybean harvest (corn–ryegrass (*Lolium multiflorum)*–soybean–winter wheat–red clover *(Trifolium pratense)*). Twenty-four soil samples were collected from the four tillage treatments in 2021.

4.3.4 Restored Prairies and Virgin Soil Sites

4.3.4.1 Restored Prairies

In 2020 and 2021 soil samples were collected at two restored prairie sites. These sites are part of the Battelle Darby Creek park which is managed by the Metro Park organization. The park features 7,196 acres of forest, prairies, and wetlands (Metroparks, n.d.). Soil samples were taken at two enclosed pastures, which are close to each other and

that are used by a bison herd throughout the seasons.

The second restored prairie site is known as the Indiana Ridge Prairies. It is a restored prairie on which purple coneflower, royal catchfly, prairie dock, big bluestem and other wildflowers and grasses are growing. The soils were identified as Celina silt loam (mesic Aquic Hapludalfs), Ockley silt loam (mesic Typic Hapludalfs), Kokomo silty clay loam (mesic Typic Argiaquolls), and Crosby silt loam (mesic Typic Hapludalfs) (Soil Survey Staff, 2019).

4.3.4.2 Virgin Soil Sites

Virgin Soil sites were defined as sites that have not been disturbed by anthropogenic activities since the first European settlers arrived in Ohio. The two sites that were included in this study are known as the W. Pearl King Prairie Savanna, which is managed by MetroParks, and the Smith Cemetery, which is managed by Ohio Department of Natural Resources. Both sites are nature preserves and provide habitat for native tallgrasses and oak groves. The soils were identified as Crosby-Lewisburg silt loam (mesic Aeric Epiaqualfs / shallow Aquic Hapludalfs).

4.3.5 Soil Sampling and Processing

Sampling sites were determined based on a soil heat map, which included LIDAR elevation information and information from the US Soil Survey website. An elevation heatmap was created using a 3D point cloud and mesh processing software CloudCompare. The soil map was overlayed with the elevation heatmap to identify a low and high elevation soil sampling site. On-farm soil sampling sites were selected

based on farmer survey information and the premise that the soil units would be identical. The GPS coordinates were recorded, and the texture specific information was obtained from the US Soil Survey website. For the restored prairies and virgin sites, the same procedure was carried out.

Six to eight soil (0-15 cm depth) cores (2.54 cm dia.) were taken and homogenized to form a composite sample (~1 kg). All cores were taken within a 5 m radius. For the Ohio LTES a randomized soil core sampling was done in a w-shaped pattern. For the Michigan LTES it was required to sample five predetermined soil sampling subplots. At each subplot two cores (0-15 cm depth; 2.54 cm dia.) were collected and composited. Soil samples were stored as soon as possible in a cooler with ice and transferred to a -20 °C freezer (Lee et al., 2007; Veum, 2019).

After thawing the soil samples in the 4 °C fridge, the wet soil was sieved to pass a 2 mm mesh size and all organic material, or mineral fragments were removed. A 300 to 500g subsample was air dried for 24 to 48 hours at room temperature, then stored in the 4 °C fridge and used to measure pH, Total C (TC), Total N (TC), soil organic carbon (SOC), and the enzyme activity of β-Glucosidase (GLU), N-Acetyl Glutamate synthase (NAG), and Arylsulfatase (AS). The remaining field moist subsample was stored at -20 ˚C and used for EL-FAME analysis work (Suppl. Fig. 7.). Gravimetric water content was determined by weighing before and after a placing a soil subsample in an oven set at 105 ˚C for 24 hours.

4.3.6 Total Nitrogen, Total Carbon, Soil Organic Carbon, and pH

Soil pH was measured with air-dried soils using a 1:1 mixture of soil and deionized water followed by measurement with a glass membrane electrode (Accumet Model 15 pH meter).

Total nitrogen (TN) and total carbon TC was determined on sieved air-dried soil samples that had been crushed with a pestle and mortar to pass a 106 μm sieve (USA Standard Test Sieve Number 104). This subgroup was then used in an elemental analyzer system (Carlo Erba CHN EA 1108, now Thermo Fisher Scientific, Waltham, MA) (Nelson & Sommers, 1996, Matejovic, 1997).

Inorganic carbon (SIC) was determined by placing the half of the subsample into a furnace for 16 hours at 450 ˚C (Ball, 1964; Davies, 1974; Ben-Door & Banin, 1989; Soon and Abboud, 1991; Nelson & Sommers, 1996). Past publications determined that organic matter content by loss-on-ignition at this 400 ˚C temperature resulted in a strong correlation with soil organic carbon content that was determined via wet-oxidation (dichromate) (Ben-Door & Banin, 1989, Nelson & Sommers, 1996). The heating regime of 375 ˚C to 450 ˚C oxidizes all organic matter without creating significant errors due to losses by crystal water or hydroxyl groups from minerals (Davies, 1974; Nelson & Sommers, 1996). After the furnace treatment the subsamples were dry combusted a second time in the elemental analyzer system. SOC was calculated by subtracting the recorded SIC concentration from the TC concentration. In the final step, TN, TC, and the SOC variable, were divided and multiplied by the percentage of clay and separately by the percentage of sand.

4.3.7 EL-FAME

The soil microbial community composition was obtained by running the Ester-Linked Fatty Acid Methyl Ester method (EL-FAME) which was described by Schutter and Dick (2000) and is based on a method developed by Dr. Rhae Drijber.

Three g of field moist soil was extracted with a 1:1 hexane/methyl-tert butyl ether and Methyl Nonadecanoate mixture that was then vortexed with a 0.2 M methanolic KOH solution. The tube was placed into a water bath at 37 ˚C and incubated for 1h. During this incubation phase the sample was vortexed for 10 seconds every 10 minutes. Afterwards 1.0 M acetic acid is added to establish a pH of 7. In the next step, 10 ml of hexane is added, and the tube is vortex for 60 seconds followed by centrifuging (1600 rpm for 20 minutes) that partitioned the EL-FAMEs were into the organic phase. The upper, organic phase was removed and evaporated under a stream of N_2 gas. The dried EL-FAME film was dissolved in 1 ml of the internal standard mixture and transferred into a gas chromatograph (GC) for analysis on the 6890N GC (Agilent Technologies).

The GC was equipped with a flame ionization detector that used a fused silica capillary column (25 m × 0.20 mm × 0.33 μm). The system used ultra-high purity H_2 as the carrier gas and the temperature program was ramped from 190 to 285 ˚C at 10 ˚C per minute. The Microbial ID PLFA identification software (MIDI ver.6.2) was used to identify the biomarker and their relative peak areas. The individual biomarkers concentrations (nmol g^{-1} dry soil) were calculated and categorized based on described procedures in the literature (Olsson, et al., 1995; Frostegård & Bååth, 1996; Zelles, 1999; Schutter and Dick, 2002).

Each EL-FAME is described with a nomenclature. The first number clarifies the number of carbon atoms of the fatty acid molecules. It is followed by a colon and a second number which explains the number of double bonds within the molecule. The suffixes "*c*" and "*t*" are used to indicate *Cis* and *trans* isomers. Branched fatty acids are indicated by the prefixes *i* (iso) and *a* (anteiso). Other notations like "*Me*", "*OH*", "*cy*" are used to describe methyl, hydroxy, and cyclopropane groups.

The total FAME concentration (nmol g^{-1} dry soil) was determined by the sum of all identified EL-FAME biomarkers in a soil sample. The sums of individual EL-FAME biomarkers were used to compute broad taxonomic microbial groups such as Gram-positive bacteria (*a*15:0, *i*15:0, *i*16:0, *a*17:0, *i*17:0) (O'Leary and Wilkinson, 1988; Zelles, 1999), Gram-negative bacteria (cy17:0, cy19:0, 16:1ω7c, 17:1ω8c, 18:1ω7c) (Wilkinson, 1988; Tunlid et al., 1989; Kerger, et al., 1986; Haack, et al., 1994, Zelles, 1999), Actinobacteria (10Me16:0, 10Me17:0, 10Me18:0, 10Me19:1ω7c) (Fischer et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Frostegård et al., 1993, Veum et al. 2021), arbuscular mycorrhizal fungi (AMF; 16:1 ω5c) (Nordby et al., 1981; Olsson et al., 1995; Olsson, 1999; Madan et al., 2002), Protozoa (20:3ω6c, 20:4ω6c) (Guckert et al., 1985), and Eukaryotes (21:0, 22:0, 23:0, and 24:0) (Zelles, 1999) (Appendix Table 2).

Additionally soil microbial ratios were calculated, which included the total fungal/bacterial ratio (tFU/BA), fungal/bacterial ration (FU/BA), gram-positive bacteria/gram-negative bacteria ratio (GP/GN), saturated/monounsaturated fatty acid ratio (SAT/MONO), bacterial/total FAME (BA/ToF), cyclopropane fatty acid 17/ 16:1 precursor ratio (Cy17/16; cy17:0/16:1ω7c), and cyclopropane fatty acid 19/ 18:1

precursor ratio (Cy19/18; cy19:0/18:1ω7c). In published studies these ratios were used to interpret microbial community shifts due to stress related conditions (McKinley et al., 2005; Taguchi et al., 1980; Guckert et al., 1986; Kieft et al., 1994, Bossio and Scow, 1998; Moore-Kucera and Dick, 2007).

The tFU/BA ratio was determined with the sum of the saprotrophic fungal and the arbuscular mycorrhizal fungi (AMF) marker (18:1ω9c, 18:2ω6c, and 16:1ω5c) divided by the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c). The FU/BA ratio is calculated in a very similar way with the exception that the AMF biomarker is removed (Frostegård & Bååth, 1996). The GP/GN ratio is calculated with the sum of 5 Gram-positive bacteria divided by 5 Gram-negative bacteria (Frostegård et al., 1993; Zelles et al., 1994). The SAT/MONO ratio was calculated with the sum of 5 saturated fatty acids (14:0, 15:0, 16:0, 17:0, and 18:0) divided by the sum of 7 monounsaturated fatty acids (16:1ω5c, 16:1ω7c, 17:1ω8c, 18:1ω7c, 18:1ω9c, cy17:0, and cy19:0) (McKinley et al., 2005). To determine the BA/ToF ratio the sum of 11 bacterial markers (15:0, 17:0, i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7c, and 18:1ω7c) was divided by the total FAME concentration. For the GP/ToF ratio the sum of 5 Gram-positive bacteria was divided by the total FAME concentration. In the final step, all recorded EL-FAME variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand to determine a possible relationship between soybean yields.

4.3.8 Enzyme Activity

The potential enzyme activity of β-glucosidase (B-GLU), Arylsulfatase (AS), and N-Acetyl-β-glutamate synthase (NAG; also known as β-glucosaminidase) were measured for each dry soil sample. These three enzyme activities are involved in the C cycle (B-GLU, NAG), S cycle (AS), and N cycle (NAG) and were determined by conducting well known enzyme assays. The assay procedures have been described elsewhere: B-GLU (Tabatabai, 1994; Dick, 2011), NAG (Parham and Deng, 2000; Dick, 2011), and AS (Tabatabai, 1994; Dick, 2011). For each enzyme assay three replicate samples and one control of 1g of air-dried soil was prepared. Each sample received the corresponding substrate based on the assay protocol before the 1-hour incubation at 37 ˚C started (Table 2.3). For the control the corresponding substrate was added after the reaction was stopped. Enzyme activities are expressed as mg of p-nitrophenol (PNP) kg^{-1} dry soil h^{-1}.

Additionally, the sum of B-GLU + AS, B-GLU + NAG, AS + NAG, and B-GLU +AS + NAG was determined. Recent studies determined that multi-assay combinations of enzymes are possible and thereby could be used as a new soil health assessment tool across agroecosystems (Acosta-Martínez et al., 2019). Additionally, the ratio of B-GLU / AS, B-GLU / NAG, AS / NAG, (B-GLU +AS) / NAG, (B-GLU + NAG) / AS, and (AS + NAG) / B-GLU were determined. In the final step, all recorded enzyme variables, were divided and multiplied by the percentage of clay and additional by the percentage of sand to determine a possible relationship between soybean yields.

4.3.9 Statistical Analysis

Statistical analyses were performed using RStudio, which is an integrated development environment that uses the R programming language and software environment for statistical computing and graphics (R Core Team, 2022; RStudio, 2018). To determine the weighted biochemical SHS variables the computational algorithm was used which is based on the optimized soybean yield model described in section 3.4.2. A list of 77 variables and their corresponding regression coefficients was used in this algorithm (Suppl. Table 4). The specific regression coefficients to calculate the SH scores in this case belonged to enzyme assay, enzyme ratio, and EL-FAMA biomarker variables.

In the next step 521 variables (included SHS) were statistically analyzed and scored for their ability to detect soil/crop management effects using a multi-level sensitivity test (Figure 4.1). Such a test does not exist in statistics and the term is used to describe the statistical analysis process. In this multi-level sensitivity test the complexity of categories is systematically increasing. In the first statistical runs each of the 521 variable is analyzed and scored for its ability to detect crop/soil management effects for cover cropping, manure amendments, and crop rotation by running t-tests ($\alpha = 0.05$). Additionally, Tukey-Kramer post hoc tests ($\alpha = 0.05$) are run for all variables where agricultural crop rotation systems are compared with natural ecosystems, and separately for the three tillage systems. These results are again scored and analyzed. These tests complete the first level of the multi-level sensitivity test. In the remaining tests the level of agricultural subgroups is increased each time by a crop/soil management system. Starting with tillage and crop rotation (e.g. corn-soy (CS), corn-

soy-wheat (CSW)), followed by cover cropping (binary numerical response of either yes or no) and last by manure (binary numerical response of either yes or no). At each level the post hoc results are analyzed and scored for all 521 variables (Figure 4.1). The same statistical analysis tests were carried for each of the three growing seasons. The process of determining the sensitivity score was described in section 3.3.11.

In total 16672 sensitivity scores were determined (521 variables * 8 categorical variable combinations * 4 years). In the next step an elimination process was used to identify the most temporarily robust, and sensitive SH indicators. The relationship between those identified SH indicators and SOC was determined and those indicators with correlation coefficients lower than 0.5 or greater than -0.5 were removed.

Furthermore, the identified SH indicators were tested for their sensitivity to detect cover cropping, manure amendments, crop rotation, and tillage effects under three SOC ranges low (<10.1 g kg^{-1}), medium (10.1 g kg^{-1} ≥ 16.2 g kg^{-1}), and high (>16.2 g kg^{-1}).

In the final statistical analysis, the identified SH indicators were used to identify detrimental and beneficial soil management practices.

4.4 Results and Discussion

4.4.1 *Identification of the most reliable and sensitive SH indicators*

The multi-level sensitivity analysis of 521 variables identified 30 temporarily robust and sensitive SH indicators (Table 4.1). The number was further reduced, to only include the 8 indicators that were significantly correlated with SOC, having r value

greater than 0.5 (Figure 4.3, Figure 4.4, Table 4.2, Table 4.3). Three of the 8 SH indicators were the EL-FAME biomarkers (total bacteria, (r = 0.57); 18:0 10-ME, (r = 0.55); 16:0 iso, (r = 0.51)). The remaining 5 indicators were SH score variables that were calculated based on the corresponding computational algorithm SHS (negative) [FAME], (r = 0.54); SHS (positive) [FAME], (r = 0.50); SHS (fitted) [FAME], (r = 0.52); SHS (fitted) [Enz+FAME], (r = 0.57); or SHS (positive) [Enz+FAME], (r = 0.54)).

4.4.2 Agricultural Crop Rotation and Natural Soil Ecosystems

The statistical analysis found that the EL-FAME total bacteria, 16:0 iso biomarker, SHS (fitted) [Enz+FAME], and SHS (positive) [Enz+FAME] variables followed the order VS=CRP>P>CSW>CS>CSW-Michigan. For the 18:0 10-ME biomarker the CRP site resulted in the highest concentrations, and the Ohio CS sites were determined to be equal to the CSW-Michigan sites resulting in the following order CRP >VS>P>CSW>CS=CSW-Michigan. For the SHS (-) [FAME] variable it was determined that P and CRP were equal in their SH scores resulting in the following order VS>CRP=P>CSW>CS>CSW-Michigan. For the SHS (+) [FAME] variable the Ohio CS sites were determined to be equal to the CSW-Michigan, and the CRP sites could not clearly be separated from the P or the VS sites resulting in the following order VS ≥CRP ≥P>CSW>CS=CSW-Michigan. If we assume that the CSW-Michigan sites resulted in the overall lowest EL-FAME concentrations and SH scores the average relative magnitude would be 4.04: 3.95: 3.08: 1.56: 1.32: 1 (VS : CRP : P : CSW : CS :

CSW-Michigan). Overall, it can be concluded that the CRP sites were identical to virgin soils and restored prairies were significantly higher than agricultural fields but significantly lower than CRP and virgin soils. CSW sites were statistically greater than Ohio and Michigan CS sites. Michigan CS sites were the lowest.

These results provide evidence that the 1985 USDA Farm bill that created the conservation reserve program (CRP) was successful in restoring agricultural soil to SH levels equal to virgin soils. Furthermore, the restored prairies resulted in EL-FAME concentrations and SH scores that were about two times higher than the ones measured on Ohio farm fields with a CSW rotation. These observations provide evidence that the four principal of soil health (maximizing continues living roots, minimizing disturbance, maximizing biodiversity, and maximizing soil cover) are beneficial for the health of a soil and that it can be quantified with the eight identified SH indicators.

The observation that the Michigan soil samples had the lowest EL-FAME concentrations and SH scores suggests there is a regional effect on the microbial community that is likely related to differences in soil type and climate. These climate effects have been noted in other investigations (Acosta-Martínez et a., 2014 a, b, Roper et al., 2017). The formation factors that specifically control the evolution of a given soil, which is parent material, climate, biota, topography, and time must be considered (Buol et al., 2011). Additionally, the significant difference between Michigan and Ohio soil microbial properties provides evidence for why the existing Soil Health tests (e.g. Haney and CASH tests) may be effective in one region but not another in detecting and quantifying SH. Furthermore, it seems likely that the development of reliable and sensitive SH indicators will require vetting them within relatively uniform agro-

ecosystems with a state or across state boundaries.

4.4.3 Cover Crops

Total Bacteria, 16:0 iso, 18:0 10-ME and all eight SH score variables were significantly greater cover cropping over no-cover cropping (Table 4.5). Fig. 4.6 shows that for seven of the eight SH indicators at the medium, and high SOC range were significantly affected by cover cropping. The only exception for this was determined for the 16:0 iso variable that was not affected by cover cropping at the high SOC level.

Under low SOC levels there was a significant effect due to cover cropping on total bacteria, SHS (fitted) [Enz+FAME], and SHS (positive) [Enz+FAME] variables. The reason that the other five SH indicators were not affected could be the explained by the sample size between cover cropping (n=36) and no cover cropping (n=9) systems.

However, these results clearly suggests that cover crops have a significant positive effect on the health of a soil, which links them to three of the four soil health principles (maximizing continues living roots, maximizing biodiversity, and maximizing soil cover).

4.4.4 Manure

Like cover crops total bacteria, 16:0 iso, 18:0 10-ME and all five SH score variables were significantly greater in manure amended soils (Table 4.6). When the data was separated into the three SOC ranges, the significant effects of manure were found for seven out of the eight SH indicators (16:0 iso being the exception) only at the highest SOC

level (Figure 4.7). Only total bacteria could detect difference at all three SOC ranges. The results suggest that manure amendments may take more time of repeated applications than what occurred in this study ($\leq$ 3 years) to significantly EL-FAME biomarkers or SH scores. This is supported by Huang et al., (2022) who reported microbial shifts on due to manure amendments on a long-term field site.

4.4.5 Tillage

Surprisingly, none of the 8 SH variables could determine significant differences between the three tillage systems (Table 4.7), even when the data was separated into the three SOC ranges (Figure 4.8). These results are contradictory to many studies where NT over PT soils has higher: enzyme activities (Lorenz et al., 2020; Balota et al., 2004, 2014; Montero et al., 2004; Acosta-Martínez et al., 2003; Bergstrom et al., 2000; Frey et al., 1999; Deng and Tabatabai, 1997; Dick, 1986a; Dick, 1986b Dick, 1984; Doran, 1980) mycorrhizal fungi biomass (Drijber et al., 2000; McGonigle et al., 1999), macro aggregation (Kumar et al., 2012a; Six et al., 2000b), TC and TN concentrations (Feng et al., 2003; Hendrix et al., 1986), SOC concentrations (Kumar et al., 2012a; Dick, 1986a; Dick, 1986b) and water holding capacity (Kumar et al., 2012a, Kumar et al., 2012b). Other studies also showed changes in the composition of microbial communities in the soils due to the tillage intensity (Jackson et al., 2003; Drijber et al., 2000; Frey et al., 1999; Doran, 1980).

These divergent results from the current study could be because most of the previous research was done LTES. Each of these sites have more or less the same soil type

where each management treatment is carefully and consistently followed over long periods. This stands in contrast to the current study which was done on farmer managed fields where there is variation in the exact management, weather conditions, and soil type among the field sites (Suppl. Fig. 5).

Because of this, the tillage data was further divided into 4 subcategories to account for differences in crop rotation, cover cropping, and manure amendment (Figure 4.9). The separation of the data into 4 subgroups now showed tillage treatment effects which follow previous research results. The no-till farms fields that practiced a corn-soy-wheat rotation, planted cover crops, and applied manure had the highest EL-FAME concentrations and SH scores for all 8 SH indicators (Table 4.8, Table 4.9, Table 4.10). The lowest concentrations and scores were determined for PT and NT farm sites with a corn-soy rotation, and no cover cropping or manure.

For the PT subgroup for the total bacteria, SHS (fitted) [Enz+FAME], and SHS (positive) [Enz+FAME] variable a significant effect was found when a corn-soy-wheat crop rotation was practiced (Table 4.8, Table 4.10). The next significant difference of PT could be detected when cover crops and manure were used resulting in the following order of: PT [CS-CCno-Mno < CSW-CCno-Mno = CSW-CCno-Myes = CSW-CCyes-Mno < CSW-CCyes-Myes]. Furthermore, the trajectory in Figure 4.9 implied that farms with intensive tillage seem to reach a SH plateau where adopting more practices that should improve soil health did not happen under this regime. The remaining 5 SH indicators could not differentiate between management systems.

For the MT subgroup 5 SH indicators (total bacteria, 18:0 10-ME, SHS (negative) [FAME], SHS (fitted) [Enz+FAME], and SHS (positive) [Enz+FAME]) lowest

EL-FAME concentrations and SH scores were the corn-soy rotation, no cover crops and manure amendment (Table 4.8, Table 4.9, Table 4.10). This observation revealed that manure amendments most likely only provided an improvement for the health of a soil under ideal conditions or long-term application if we assume that the 4-Rs for nutrient management were followed (Right Source, Right Rate, Right Time, and Right Place). The next significant difference in the MT subgroups was found when cover crops were used. After a CSW crop rotation was the only management practice that resulted in significantly greater concentration and scores. This observation provided evidence that the MT similar to the PT subgroup was reaching its SH plateau, resulting in a final order of: MT [CS-CCno-Myes <= CS-CCno-Mno <= CS-CCyes-Mno < CSW-CCno-Mno >= CSW-CCno-Myes = CSW-CCyes-Myes].

For NT all 5 SH score variables were determined to be the most sensitive variables (Table 4.9, Table 4.10). The soil health scores showed a small increase for CS farm sites that used manure. But the statistical increase was significantly higher for CS farm sites that used cover crops. CS farm sites that used manure and cover crops matched those that followed a CSW crop rotation with cover cropping and no manure amendment. Overall, the NT subgroup was the only one that displayed a steady increase with the introduction of soil amendments, and biodiversity maximization resulting in this final order of: NT [CS-CCno-Mno <= CS-CCno-Myes <= CS-CCyes-Mno <= CS-CCyes-Myes = CSW-CCyes-Mno < CSW-CCyes-Myes].

Overall, the data analysis suggested that regardless of tillage system, a CSW rotation was preferable for increasing SH, followed by cover cropping, and manure amendments having the smallest effect on SH. The statistical analysis also found that a PT

in combination with CSW- CCyes-Myes had a similar SH score to MT with CSW-CCno-Myes and NT with CS-CCyes-Mno (Figure 4.9). Furthermore, the four principles of SH are met by the management system that received a high SH score.

4.4.6 Interpreting the optimized Multivariate Soybean prediction model

A simulation modeling experiment was done using the optimized soybean prediction model (Table 4.11). The objective was to determine the highest theoretical soybean yield. This model assumes the following conditions were met: (1) conventional management, (2) corn-soy-wheat rotation, (3) applied cattle manure once every 3 years, (4) cover crops grown once every 3 years, (5) chisel tilled (MT), (6) a pH 6 to 7, (7) planting in April, (8) weed control with glyphosate, (9) no fungicide, (10) a high soil surface residue coverage. Additionally, it was assumed that the soil was a Mollisol, and it has the highest recorded SOC and TN concentrations, and it has a Soil Health Score of 1. With these assumptions a theoretical soybean yield of 16025 kg/ha (238 bsh/ac) would be possible. Surprisingly, this number came close to the soybean world record of 206 bsh/ac grown in 2019 by Randy Dowdy in Georgia (Bennet, 2023).

On the other hand, the model predicted that if the following conditions were met for a conventional managed field with: (1) corn-soy rotation, (2) no manure amendments, (3) no cover cropping, (4) intensive tillage (PT), (5) pH ≥ 8, (6) planting in May, (7) weed control with glyphosate dicamba, (8) fungicide treatment, (9) residue cover below 6 %, and (10) soil health scores at 0.2 the theoretically soybean yield would be at 0 kg/ha.

This exercise demonstrates that these models are often only approximations

and that predicted yields might not occur for given field or cropping year – due extraneous factors such as extreme weather, pest/disease infestations or negative management for which SH has limited control to mitigate. However, models showed that biological soil properties are important or related to soil health and sustainable soil management.

The simulated experiment showed that the SHS was a significant factor in predicting yield. The SHS influenced 53 to 88% of the theoretical soybean yields (Figure 4.10). Furthermore, this connects healthy soils to optimized agriculture productivity.

The model showed the importance of soil texture for soybean yields. The information from the optimized model predicted that 50 kg/ha (0.74 bsh/ac) would be lost for each percentage point reduction in clay content (Figure 4.11).

4.5 Conclusions

A multi-level sensitivity assessment of 521 soil properties, under various agricultural and non-agricultural management systems, identified 30 robust and sensitive Soil Health (SH) indicators. Selection of indicators that had a significant correlation with SOC reduced the number of sensitive SH indicators to eight. The statistical analysis of the final 8 SH indicators showed that they are significantly influenced by regional conditions, and that the comparison between agricultural and non-agricultural results confirm the four principal of soil health (maximizing continues living roots, minimizing disturbance, maximizing biodiversity, and maximizing soil cover). This study, therefore, represents an important role that microbial and

enzymatic ecology components for detecting and quantifying management effects on SH.

A separate analysis of the 8 SH indicators focused on agricultural management systems in Ohio, showed that corn-soybean-wheat crop rotation, cover cropping, and manure had a significantly effect on all 8 SH indicators.

Pairwise analysis of tillage systems revealed no significant effects for all 8 SH indicators. The analysis of tillage systems under the low, medium, or high SOC levels did not change this result. Because of the significant effects of crop rotation, cover crops and manure, a follow-up analysis was conducted. Tillage systems were divided into subcategories that were based on other land management practices. The statistical subcategory analysis allowed the identification of detrimental or beneficial agricultural management systems for the health of a soil. It confirmed that the four principles of SH were applicable at the farm field scales and suggested that intensive tillage and minimal tillage there is no adding more than one management system that should improve SH did not happen. The post-hoc Tukey was used to rank the 8 SH indicators for tillage regime from the most sensitive to least sensitive with a final order of: SHS (fitted) [Enz+FAME], SHS (positive) [Enz+FAME], total bacteria, SHS (negative) [FAME], 18:0 10-ME, SHS (positive) [FAME], SHS (fitted) [FAME], 16:0 iso.

4.6 References

Acosta-Martínez, V., Zobeck, T. M., Gill, T. E., & Kennedy, A. C., 2003. Enzyme activities and microbial community structure in semiarid agricultural soils. Biology and Fertility of Soils, 38(4), 216–227. https://doi.org/10.1007/s00374-003-0626-1

Acosta-Martínez, V., Cruz, L., Sotomayor-Ramírez, D., & Pérez-Alegría, L., 2007a. Enzyme activities as affected by soil properties and land use in a tropical watershed. Applied Soil Ecology, 35(1), 35–45. https://doi.org/10.1016/j.apsoil.2006.05.012

Acosta-Martínez, V., Mikha, M. M., & Vigil, M. F., 2007b. Microbial communities and enzyme activities in soils under alternative crop rotations compared to wheat–fallow for the Central Great Plains. Applied Soil Ecology, 37(1-2), 41–52. https://doi.org/10.1016/j.apsoil.2007.03.009

Acosta-Martínez, V., Cotton, J., Gardner, T., Moore-Kucera, J., Zak, J., Wester, D., & Cox, S., 2014a. Predominant bacterial and fungal assemblages in agricultural soils during a record drought/heat wave and linkages to enzyme activities of biogeochemical cycling. Applied Soil Ecology, 84, 69–82. https://doi.org/10.1016/j.apsoil.2014.06.005

Acosta-Martinez, V., Moore-Kucera, J., Cotton, J., Gardner, T., & Wester, D., 2014b. Soil enzyme activities during the 2011 Texas record drought/heat wave and implications to biogeochemical cycling and organic matter dynamics. Applied Soil Ecology, 75, 43–51. https://doi.org/10.1016/j.apsoil.2013.10.008

Acosta-Martínez, V., Pérez-Guzmán, L., & Johnson, J. M. F., 2019. Simultaneous determination of β-glucosidase, β-glucosaminidase, acid phosphomonoesterase, and arylsulfatase activities in a soil sample for a biogeochemical cycling index. Applied Soil Ecology, 142, 72–80. https://doi.org/10.1016/j.apsoil.2019.05.001

Ball, D.F., 1964. Loss-on-Ignition as an estimate of organic matter and organic carbon ion non-calcareous soils. Journal of Soil Science, 15: 84-92. https://doi-org.proxy.lib.ohio-state.edu/10.1111/j.1365-2389.1964.tb00247.x

Bandick, A. K., & Dick, R. P., 1999. Field management effects on soil enzyme activities. Soil Biology and Biochemistry, 31(11), 1471–1479. https://doi.org/10.1016/s0038-0717(99)00051-6

Ben-Dor E. and Banin A., 1989. Determination of organic matter content in arid-zone soils using a simple “loss-on-ignition” method, Communications in Soil Science and Plant Analysis, 20:15-16, 1675 1695, DOI: 10.1080/00103628909368175

Bennett, C., 2023. Young Farmer Breaks Soybean World Record With Stunning 206-Bushel Yield. AGWEB. https://www.agweb.com/news/crops/soybeans/young-farmer-breaks-soybean-world-record-stunning-206-bushel-yield

Bergstrom, D. W., Monreal, C. M., Tomlin, A. D., & Miller, J. J., 2000. Interpretation of soil enzyme activities in a comparison of tillage practices along a topographic and textural gradient. Canadian Journal of Soil Science, 80(1), 71–79. https://doi.org/10.4141/s99-034

Blanco-Canqui, H., Lal, R., 2007. Soil structure and organic carbon relationships following 10 years of wheat straw management in no-till, Soil and Tillage Research, Volume 95, Issues 1–2, Pages 240-254, https://doi.org/10.1016/j.still.2007.01.004

Bossio, D., Scow, K., 1998. Impacts of Carbon and Flooding on Soil Microbial Communities: Phospholipid Fatty Acid Profiles and Substrate Utilization Patterns. Microb Ecol 35, 265–278. https://doi.org/10.1007/s002489900082

Bowman, J. P., Skerratt, J. H., Nichols, P. D., & Sly, L. I., 1991. Phospholipid fatty acid and lipopolysaccharide fatty acid signature lipids in methane-utilizing bacteria. FEMS Microbiology Letters, 85(1), 15–22. https://doi.org/10.1111/j.1574-6968.1991.tb04693.x

Bowman, J. P., Sly, L. I., Nichols, P. D., & Hayward, A. C., 1993. Revised Taxonomy of the Methanotrophs: Description of Methylobacter gen. nov., Emendation of Methylococcus, Validation of Methylosinus and Methylocystis Species, and a Proposal that the Family Methylococcaceae Includes Only the Group I Methanotrophs. International Journal of Systematic Bacteriology, 43(4), 735–753. https://doi.org/10.1099/00207713-43-4-735

Buol, S., 2013. Soil genesis and classification. CATENA. https://www.academia.edu/34154456/Soil_genesis_and_classification

Carlson, J., Saxena, J., Basta, N. et al., 2015. Application of organic amendments to restore degraded soil: effects on soil microbial properties. Environ Monit Assess 187, 109. https://doi.org/10.1007/s10661-015-4293-0

Chahal, I., & Van Eerd, L. L. (2019). Quantifying soil quality in a horticultural-cover cropping system. Geoderma, 352, 38–48. https://doi.org/10.1016/j.geoderma.2019.05.039

Chu, M., Singh, S., Walker, F. R., Eash, N. S., Buschermohle, M. J., Duncan, L. A., & Jagadamma, S. (2019). Soil Health and Soil Fertility Assessment by the Haney Soil Health Test in an Agricultural Soil in West Tennessee. Communications in Soil Science and Plant Analysis, 50(9), 1123–1131. https://doi.org/10.1080/00103624.2019.1604731

Climate Fieldview. (n.d.). [Extracting precipitation amounts for sites located in Ohio and Michigan]. Retrieved between September 2019 until December 2021

Davies, B. E., 1974. Loss-on-Ignition as an Estimate of Soil Organic Matter1. Soil Science Society of America Journal, 38(1), 150. https://doi.org/10.2136/sssaj1974.03615995003800010046x

Deiss, L., Sall, A., Demyan, M. S., & Culman, S. W., 2021. Does crop rotation affect soil organic matter stratification in tillage systems? Soil and Tillage Research, 209, 104932. https://doi.org/10.1016/j.still.2021.104932

Deng, S. P., & Tabatabai, M. A., 1997. Effect of tillage and residue management on enzyme activities in soils: III. Phosphatases and arylsulfatase. Biology and Fertility of Soils, 24(2), 141–146. https://doi.org/10.1007/s003740050222

Dick, R.P., Rasmussen P.E., & Kerle E.A., 1988. Influence of long-term residue management on soil enzyme activity in relation to soil chemical properties of a wheat-fallow system. Biol. Fert. Soils. 6:159-164.

Dick, R. P., 2011. Methods of Soil Enzymology. In Soil Science Society of America book series. https://doi.org/10.2136/sssabookser9Dick, W. A., 1984. Influence of Long-Term Tillage and Crop Rotation Combinations on Soil Enzyme Activities. Soil Science Society of America Journal, 48(3), 569–574. https://doi.org/10.2136/sssaj1984.03615995004800030020x

Dick, W. A., 1984. Influence of Long-Term Tillage and Crop Rotation Combinations on Soil Enzyme Activities. Soil Science Society of America Journal, 48(3), 569–574. https://doi.org/10.2136/sssaj1984.03615995004800030020x

Dick, W.A. and D.M. Van Doren, Jr., 1985. Continuous tillage and rotation combinations effects on corn, soybean, and oat yields. Agron. J. 77:459-465.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986a. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Typic Fragiudulf Soil. The Ohio State University, Ohio Agricultural Research and Development Center Wooster, Ohio, p. 34. Research Bulletin ll81.

Dick, W.A., Van Doren Jr., D.M., Triplett, G.B., Henry, J.E., 1986b. Influence of long-term tillage and rotation combinations on crop yields and selected soil parameters. Results Obtained for a Mollic Ochraqualf Soil. Research Bulletin 1180. The Ohio State University Ohio Agricultural Research and Development Center, Wooster, Ohio, 30 p.

Dick, W.A., Triplett, G.B., Van Doren Jr., D.M., 2013. Triplett-Van Doren Long-term Tillage and Crop Rotation Data (Accessed 15 December 2019). Ohio Agricultural Research and Development Center, Wooster, Ohio. https://kb.osu.edu/handle/1811/55716

Doran, J. W., 1980. Microbial Changes Associated with Residue Management with Reduced Tillage. Soil Science Society of America Journal, 44(3), 518–524. https://doi.org/10.2136/sssaj1980.03615995004400030016x

Dowling, N.J.E., Nichols, P.D. and White, D.C., 1988. Phospholipid fatty acid and infra-red spectroscopic analysis of a sulfate-reducing consortium. FEMS Microbiol. Ecol. 53, 325-334.

Dowling, N.J.E., Widdel, F. and White, D.C., 1986. Phospholipid Ester-linked Fatty Acid Biomarkers of Acetate-oxidizing Sulphate-reducers and Other Sulphide-forming Bacteria J. Gen. Microbial. 132, 1815-1829

Drijber, R. A., Doran, J. W., Parkhurst, A. M., & Lyon, D. J., 2000. Changes in soil microbial community structure with tillage under long-term wheat-fallow management. Soil Biology and Biochemistry, 32(10), 1419–1430. https://doi.org/10.1016/s0038-0717(00)00060-2

Faé, G. S., Kemanian, A. R., Roth, G. W., White, C., & Watson, J. E., 2020. Soybean yield in relation to environmental and soil properties. European Journal of Agronomy, 118, 126070. https://doi.org/10.1016/j.eja.2020.126070

Federle, T. W. 1986. Microbial distribution in soil-new techniques, p. 493-498. In F. Megusar and M. Gantar (ed.), Perspectives in microbial ecology. Slovene Society for Microbiology, Ljubljana, Slovenia.

Feng, Y., Motta, A. C., Reeves, D. W., Burmester, C. H., van Santen, E., & Osborne, J. A., 2003. Soil microbial communities under conventional-till and no-till continuous cotton systems. Soil Biology and Biochemistry, 35(12), 1693–1703. https://doi.org/10.1016/j.soilbio.2003.08.016

Fischer, A., Kroppenstedt, R. M., & Stackebrandt, E., 1983. Molecular-genetic and Chemotaxonomic Studies on Actinomadura and Nocardiopsis. Microbiology, 129(11), 3433–3446. https://doi.org/10.1099/00221287-129-11-3433

Frey, S. D., Elliott, E. T., & Paustian, K., 1999. Bacterial and fungal abundance and biomass in conventional and no-tillage agroecosystems along two climatic gradients. Soil Biology and Biochemistry, 31(4), 573–585. https://doi.org/10.1016/s0038-0717(98)00161-8

Frostegård, Å., Tunlid, A., & Bååth, E., 1993. Phospholipid Fatty Acid Composition, Biomass, and Activity of Microbial Communities from Two Soil Types Experimentally Exposed to Different Heavy Metals. Applied and Environmental Microbiology, 59(11), 3605–3617. https://doi.org/10.1128/aem.59.11.3605-3617.1993

Frostegård, Å., & Bååth, E., 1996. The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. Biology and Fertility of Soils, 22(1-2), 59–65. https://doi.org/10.1007/bf00384433

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D., 1985. Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. Fems Microbiology Letters, 31(3), 147–158. https://doi.org/10.1111/j.1574-6968.1985.tb01143.x

Haack, S.K., Garchow, H., Odelson, D.A., Forney, L.J., Klug, M.J., 1994. Accuracy, reproducibility, and interpretation of fatty acid methyl ester profiles of model bacterial communities. Appl. Environ. Microbiol. 60, 2483-2493.

Hendrix, P. F., Parmelee, R. W., Crossley, D. A., Coleman, D. C., Odum, E. P., & Groffman, P. M., 1986. Detritus Food Webs in Conventional and No-Tillage Agroecosystems. BioScience, 36(6), 374–380. https://doi.org/10.2307/1310259

Huang, X., Jia, Z., Jiao, X., Wang, J., & Huang, X., 2022. Long-term manure applications to increase carbon sequestration and macroaggregate-stabilized carbon. Soil Biology and Biochemistry, 174, 108827. https://doi.org/10.1016/j.soilbio.2022.108827

Jackson, L. E., Calderon, F. J., Steenwerth, K. L., Scow, K. M., & Rolston, D. E., 2003. Responses of soil microbial processes and community structure to tillage events and implications for soil quality. Geoderma, 114(3-4), 305–317. https://doi.org/10.1016/s0016-7061(03)00046-6

Kerger, B. D., Nichols, P. D., Antworth, C. P., Sand, W., Bock, E., Cox, J. C., Langworthy, T. A., & White, D. C., 1986. Signature fatty acids in the polar lipids of acid-producing Thiobacillus spp.: Methoxy, cyclopropyl, alpha-hydroxy-cyclopropyl and branched and normal monoenoic fatty acids. FEMS Microbiology Letters, 38(2), 67–77. https://doi.org/10.1016/0378-1097(86)90144-8

Kieft,T.L., Ringelberg D.B., White D.C..1994.Changesinester-linkedphospholipid fatty acid profiles of subsurface bacteria during starvation anddesiccation in porous medium. Appl. Environ. Microbiol.60:3292–3299. DOI: https://doi.org/10.1128/aem.60.9.3292-3299.1994

Kieft, T. L., Wilch, E., O'connor, K., Ringelberg, D. B., & White, D. C., 1997. Survival and phospholipid Fatty Acid profiles of surface and subsurface bacteria in natural sediment microcosms. Applied and Environmental Microbiology, 63(4), 1531–1542. https://doi.org/10.1128/aem.63.4.1531-1542.1997

Klose, S., Moore, J. M., & Tabatabai, M. A., 1999. Arylsulfatase activity of microbial biomass in soils as affected by cropping systems. Biology and Fertility of Soils, 29(1), 46–54. https://doi.org/10.1007/s003740050523

Knight, T. R., & Dick, R. P., 2004. Differentiating microbial and stabilized β-glucosidase activity relative to soil quality. Soil Biology and Biochemistry, 36(12), 2089–2096. https://doi.org/10.1016/j.soilbio.2004.06.007

Kroppenstedt, R., 1985. Fatty acid and menaquinone analysis of actinomycetes and related organisms. In: Goodfellow, M., Minnikin, D.E. (Eds.), Chemical Methods in Bacterial Systematics. SAB Technical Series 20. Academic Press, London, United Kingdom, pp. 173–199.

Kumar, S., Kadono, A., Lal, R., & Dick, W., 2012a. Long-Term No-Till Impacts on Organic Carbon and Properties of Two Contrasting Soils and Corn Yields in Ohio. Soil Science Society of America Journal, 76(5), 1798–1809. https://doi.org/10.2136/sssaj2012.0055

Kumar, S., Kadono, A., Lal, R., & Dick, W., 2012b. Long-Term Tillage and Crop Rotations for 47–49 Years Influences Hydrological Properties of Two Soils in Ohio. Soil Science Society of America Journal, 76(6), 2195–2207. https://doi.org/10.2136/sssaj2012.0098

Lee, Y. B., 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), 1299–1305. https://doi.org/10.2136/sssaj2006.0245

Liptzin, D., Norris, C. E., Cappellazzi, S. B., Bean, G. M., Cope, M., Greub, K. L. H., Rieke, E. L., Tracy, P. W., Aberle, E., Ashworth, A., Bañuelos Tavarez, O., Bary, A. I., Baumhardt, R. L., Borbón Gracia, A., Brainard, D. C., Brennan, J. R., Briones Reyes, D., Bruhjell, D., Carlyle, C. N., & Crawford, J. J. W. (2022). An evaluation of carbon indicators of soil health in long-term agricultural experiments. Soil Biology and Biochemistry, 172, 108708. https://doi.org/10.1016/j.soilbio.2022.108708

Liptzin, D., Rieke, E. L., Cappellazzi, S. B., G. Mac Bean, Cope, M., Kelsey, et al., 2023. An evaluation of nitrogen indicators for soil health in long-term agricultural experiments. Soil Science Society of America Journal, 87(4), 868–884. https://doi.org/10.1002/saj2.20558

Lorenz, N., B. McSpadden Gardener, B., R. Lee, N., Ramsier, C., & P. Dick, R., 2020. Soil Enzyme Activities Associated with Differential Outcomes of Contrasting Approaches to Soil Fertility Management in Corn and Soybean Fields. Applied Ecology and Environmental Sciences, 8(6), 517–525. https://doi.org/10.12691/aees-8-6-26

Madan, R., Pankhurst, C., Hawke, B., & Smith, S., 2002. Use of fatty acids for identification of AM fungi and estimation of the biomass of AM spores in soil. Soil Biology and Biochemistry, 34(1), 125–128. https://doi.org/10.1016/S0038-0717(01)00151-1

Martin, T., & Sprunger, C.D., 2022. Sensitive Measures of Soil Health Reveal Carbon Stability Across a Management Intensity and Plant Biodiversity Gradient. Frontiers in Soil Science. Volume 2, doi: 10.3389/fsoil.2022.917885

Matejovic, I., 1997. Determination of carbon and nitrogen in samples of various soils by dry combustion. Communications in Soil Science and Plant Analysis 28, 1499–1511.

McGonigle, T. P., Miller, M. H., & Young, D., 1999. *Plant and Soil*, *210*(1), 33–42. https://doi.org/10.1023/a:1004633512450

McKinley, V. L., Peacock, A. D., & White, D. C., 2005. Microbial community PLFA and PHB responses to ecosystem restoration in tallgrass prairie soils. Soil Biology and Biochemistry, 37(10), 1946–1958. https://doi.org/10.1016/j.soilbio.2005.02.033

MetroParks. (n.d.). Battelle Darby Creek. https://www.metroparks.net/parks-and-trails/battelle-darby-creek/

Miller, M., & Dick, R. P., 1995. Dynamics of soil C and microbial biomass in whole soil and aggregates in two cropping systems. Applied Soil Ecology, 2(4), 253–261. https://doi.org/10.1016/0929-1393(95)00060-6

Moore-Kucera, J., & Dick, R. P., 2007. PLFA Profiling of Microbial Community Structure and Seasonal Shifts in Soils of a Douglas-fir Chronosequence. Microbial Ecology, 55(3), 500–511. https://doi.org/10.1007/s00248-007-9295-1

Nawaz, A., Lal, R., Shrestha, R. K., & Farooq, M., 2016. Mulching Affects Soil Properties and Greenhouse Gas Emissions Under Long-Term No-Till and Plough-Till Systems in Alfisol of Central Ohio. Land Degradation & Development, 28(2), 673–681. https://doi.org/10.1002/ldr.2553

Nelson, D.W. and Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter. In Methods of Soil Analysis (eds D.L. Sparks, A.L. Page, P.A. Helmke, R.H. Loeppert, P.N. Soltanpour, M.A. Tabatabai, C.T. Johnston and M.E. Sumner). https://doi.org/10.2136/sssabookser5.3.c34

Nichols, P. D., Smith Glen A, Antworth, C. P., Hanson, R. S., & White, D. C., 1985. Phospholipid and lipopolysaccharide normal and hydroxy fatty acids as potential signatures for methane-oxidizing bacteria. FEMS Microbiology Letters, 31(6), 327–335. https://doi.org/10.1111/j.1574-6968.1985.tb01168.x

Nordby, H. E., Nemec, S., & Nagy, S., 1981. Fatty acids and sterols associated with citrus root mycorrhizae. Journal of Agricultural and Food Chemistry, 29(2), 396–401. https://doi.org/10.1021/jf00104a043

Norris, C. E., Bean, G. M., Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W., 2020. Introducing the North American project to evaluate soil health measurements. Agronomy Journal, 112(4), 3195–3215. https://doi.org/10.1002/agj2.20234

O'Leary, W. M., and S. G. Wilkinson. 1988. Gram-positive bacteria, p. 117-185. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Olsson, P. A., Bååth, E., Jakobsen, I., & Söderström, B., 1995. The use of phospholipid and neutral lipid fatty acids to estimate biomass of arbuscular mycorrhizal fungi in soil. Mycological Research, 99(5), 623–629. https://doi.org/10.1016/s0953-7562(09)80723-5

Olsson, P. A., 1999. Signature fatty acids provide tools for determination of the distribution and interactions of mycorrhizal fungi in soil. 29(4), 303–310. https://doi.org/10.1111/j.1574-6941.1999.tb00621.x

Parham, J. A., & Deng, S. P., 2000. Detection, quantification and characterization of β-glucosaminidase activity in soil. Soil Biology and Biochemistry, 32(8-9), 1183–1190. https://doi.org/10.1016/s0038-0717(00)00034-1

Parker, J. H., Smith, G. A., Fredrickson, H. L., J. Robie Vestal, & White, D. C., 1982. Sensitive assay, based on hydroxy fatty acids from lipopolysaccharide lipid A, for Gram-negative bacteria in sediments. Applied and Environmental Microbiology, 44(5), 1170–1177. https://doi.org/10.1128/aem.44.5.1170-1177.1982

Parkes, R. J., Dowling, N. J. E., White, D. C., Herbert, R. A., & Gibson, G., 1993. Characterization of sulphate-reducing bacterial populations within marine and estuarine sediments with different rates of sulphate reduction. FEMS Microbiology Ecology, 102, 235–250. https://centaur.reading.ac.uk/34143/

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio., 2018. RStudio: Integrated development for R. RStudio.

Robertson, G., D. Wall, K. Gross, and K. Renner., 2020. Baseline Spatial Variability Study at the Kellogg Biological Station, Hickory Corners, MI (1988) ver 36. Environmental Data Initiative. https://doi.org/10.6073/pasta/b49a09795d29e0c1149efcc068f3426e

Robertson, G., Snapp S., 2019. Agronomic Yields in Row Crop Agriculture at the Kellogg Biological Station, Hickory Corners, MI (1989 to 2018) ver 39. Environmental Data Initiative. https://doi.org/10.6073/pasta/00c3343378bf55e0727cccecc8a8ff41

Robertson, G., Simmons J., 2020. Main Cropping System Experiment Field Logs and treatment descriptions at the Kellogg Biological Station, Hickory Corners, MI (1988 to 2020) ver 90. Environmental Data Initiative. https://doi.org/10.6073/pasta/e5642426f9200404b6ba3af45ad9c423

Roper, W.R., Osmond, D.L., Heitman, J.L., Wagger, M.G. and Reberg-Horton, S.C., 2017. Soil Health Indicators Do Not Differentiate among Agronomic Management Systems in North Carolina Soils. Soil Science Society of America Journal, 81: 828-843. https://doi.org/10.2136/sssaj2016.12.0400

Saroa, G. S., & Lal, R., 2003. Soil restorative effects of mulching on aggregation and carbon sequestration in a Miamian soil in central Ohio. Land Degradation & Development, 14(5), 481–493. https://doi.org/10.1002/ldr.569

Schutter, M. E., & Dick, R. P., 2000. Comparison of Fatty Acid Methyl Ester (FAME) Methods for Characterizing Microbial Communities. Soil Science Society of America Journal, 64(5), 1659–1668. https://doi.org/10.2136/sssaj2000.6451659x

Schutter, M. E., & Dick, R. P., 2002. Microbial Community Profiles and Activities among Aggregates of Winter Fallow and Cover-Cropped Soil. Soil Science Society of America Journal, 66(1), 142–153. https://doi.org/10.2136/sssaj2002.1420

Skujiņš, J., 1978. History of abiontic soil enzyme research. In: Burns RG (ed) Soil enzymes. Academic Press, London, pp 1-49

Six, J., Paustian, K., Elliott, E. T., & Combrink, C., 2000. Soil Structure and Organic Matter. Soil Science Society of America Journal, 64(2), 681. https://doi.org/10.2136/sssaj2000.642681x

Soil Survey Staff., 2019. Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online. Accessed [Jan/2019-2023]

Soon Y. K., Abboud S., 1991. A comparison of some methods for soil organic carbon determination, Communications in Soil Science and Plant Analysis, 22:9-10, 943-954, DOI: 10.1080/00103629109368465

Tabatabai, M.A. 1994. Soil enzymes. In: R.W. Weaver, J.S. Angle, and P.S. Bottomley, editors, Methods of soil analysis: Part 2. Microbiological and biochemical properties. Vol. 5. Madison, WI: SSSA. p. 775–833.

Taguchi, M., Izui, K., Katsuki, H., 1980. Augmentation of cyclopropane fatty acid synthesis under stringent control in Escherichia coli. J. Biochem. 88, 1879–1882.

Tunlid, A., Hoitink, H.A.J., Low, C. and White, D.C., 1989. Characterization of bacteria that suppress Rhizoctonia damping-off in bark compost media by analysis of fatty acid biomarkers. Applied and environmental microbiology, 55(6), pp.1368-1374.

Vestal, J. R., & White, D. C., 1989. Lipid Analysis in Microbial Ecology. BioScience, 39(8), 535–541. https://doi.org/10.2307/1310976

Veum, K.S., Lorenz, T. & Kremer, R.J., 2019., Phospholipid Fatty Acid Profiles of Soils under Variable Handling and Storage Conditions. Agronomy Journal, 111: 1090-1096. https://doi.org/10.2134/agronj2018.09.0628

Veum, K. S., Acosta-Martinez, V., R. Michael Lehman, Li, C., Cano, A., & Nunes, M. R., 2021. PLFA and EL-FAME Indicators of Microbial Community Composition. ASSA, CSSA and SSSA, 251–288. https://doi.org/10.1002/9780891189831.ch12

Wallis, C. M., Lawrence, D. P., Travadon, R., & Baumgartner, K., 2021. Characterization of grapevine fungal canker pathogens Fatty Acid Methyl Ester (FAME) profiles. Mycologia, 1–11. https://doi.org/10.1080/00275514.2021.1983396

Wilkinson, S. G., 1988. Gram-negative bacteria, p. 299-457. In C. Ratledge and S. G. Wilkinson (ed.), Microbial lipids, vol. 1. Academic Press Ltd., London.

Zelles, L., Bai, Q. Y., Ma, R. X., Rackwitz, R., Winter, K., & Beese, F., 1994. Microbial biomass, metabolic activity and nutritional status determined from fatty acid patterns and poly-hydroxybutyrate in agriculturally-managed soils. *Soil Biology and Biochemistry*, *26*(4), 439–446. https://doi.org/10.1016/0038-0717(94)90175-9

Zelles, L., 1997. Phospholipid fatty acid profiles in selected members of soil microbial communities. Chemosphere, 35(1-2), 275–294. https://doi.org/10.1016/s0045-6535(97)00155-0

Zelles, L., 1999. Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. Biology and Fertility of Soils, 29(2), 111–129. https://doi.org/10.1007/s003740050533

Table 4.1. 30 SH indicators identified by the multi-level sensitivity test and their sensitivity scores (max = 20).

EL-FAME	Sensitivity Score	Enzyme Assays	Sensitivity Score	SHS Variable	Sensitivity Score
Eukaryotes / Clay %	9.84	NAG / (1 - Sand %)	9.00	SHS.FAME.neg / Clay %	9.34
Total Fungi	9.25	NAG / Clay %	8.95	SHS.FAME.pos / Clay %	9.34
Gram (-) / Clay %	9.53	NAG / Sand %	8.65	SHS.FAME.fit / Clay %	9.11
22:0 / Clay %	9.46	**NAG**	8.50	SHS.Enz.FAME.neg / Clay %	8.50
18:0 10-ME / Clay %	9.04			SHS.Enz.FAME.pos / Clay %	8.41
17:0 cyclo w7c / Clay %	8.30			SHS.Enz.FAME.fit / Clay %	8.41
Total FAME / Clay %	8.60			**SHS.FAME.neg**	7.86
17:0 10-ME / Clay %	8.02			**SHS.FAME.pos**	8.09
20:4 w6c	8.16			**SHS.FAME.fit**	7.78
Gram (-) Bacteria	8.11			**SHS.Enz.FAME.fit**	6.73
Eukaryotes	7.69			**SHS.Enz.FAME.pos**	6.57
18:0 10-methyl	7.64				
Total Bacteria	7.55				
16:0 / Clay %	7.68				
16:0 iso	7.34				

Table 4.2. Correlation Matrix of 13 SH indicators identified in the multi-level sensitivity test which were used for developing differentiation scores from the LTES (Michigan LTES data was excluded).

	TN	Total Fungi	20:4 w6c	Gram (-) Bacteria	Eukary.	18:0 10-ME	Total Bacteria	16:0 iso	NAG	SHS_ FAME (-)	SHS_ FAME (+)	SHS_ FAME (fit)	SHS_ All (fit)	SHS_ All (+)
SOC	0.90***	0.15 *	0.12 n.s.	0.41***	0.37***	0.55***	0.57***	0.51***	0.26***	0.54***	0.50***	0.52***	0.57***	0.54***
TN		0.18 **	0.14 *	0.41***	0.35***	0.52***	0.56***	0.49***	0.33***	0.53***	0.49***	0.51***	0.57***	0.54***
Total Fungi			0.42***	0.81***	0.47***	0.56***	0.61***	0.57***	0.47***	0.59***	0.62***	0.61***	0.58***	0.59***
20:4 w6c				0.43***	0.15 *	0.35***	0.45***	0.32***	0.43***	0.37***	0.35***	0.36***	0.38***	0.37***
Gram (-) Bac					0.61***	0.81***	0.90***	0.84***	0.62***	0.82***	0.84***	0.83***	0.84***	0.84***
Eukaryotes						0.59***	0.56***	0.60***	0.46***	0.88***	0.90***	0.90***	0.80***	0.82***
18:0 10-ME							0.90***	0.91***	0.51***	0.82***	0.82***	0.82***	0.84***	0.84***
Total Bacteria								0.92***	0.64***	0.87***	0.86***	0.86***	0.90***	0.89***
16:0 iso									0.55***	0.85***	0.83***	0.84***	0.84***	0.83***
NAG										0.63***	0.63***	0.63***	0.68***	0.69***
SHS_FAME (-)											0.99***	1.00***	0.96***	0.96***
SHS_FAME (+)												1.00***	0.96***	0.97***
SHS_FAME (fit)													0.96***	0.97***
SHS_All (fit)														1.00***

*** Signifies a model with a P-value <0.001.
** Signifies a model with a P-value <0.01.
* Signifies a model with a P-value <0.05.
n.s. not significant (P > 0.05)

Table 4.3. 30 Soil Health indicators organized by their ability to differentiate between agricultural management systems and their corresponding differentiation scores determined in Chapter 3. Highlighted variables in green were fair correlations (r >= 0.5) with soil organic carbon (SOC) and represent the most reliable SH indictors (Figure.4.3, Table 4.4).

High Potential	Diff. Score [LTES]	Some Potential	Diff. Score [LTES]	More Data required	Diff. Score [LTES]
Total Fungi	2.25	NAG	1.5	Eukaryotes / Clay %	N/A
Eukaryotes	2.75	20:4 w6c	1.5	Gram(-) / Clay %	N/A
Total Bacteria	2.25	Gram (-) Bacteria	1.5	22:0 / Clay %	N/A
16:0 iso	3.25	**18:0 10-ME**	1.75	18:0 10-ME / Clay %	N/A
SHS.FAME.neg	2.5			17:0 cyclo w7c / Clay %	N/A
SHS.FAME.pos	2.5			Total FAME / Clay %	N/A
SHS.FAME.fit	2.5			17:0 10-ME / Clay %	N/A
SHS.Enz.FAME.fit	3.25			16:0 / Clay %	N/A
SHS.Enz.FAME.pos	2.75			NAG / (1 - Sand %)	N/A
				NAG / Clay %	N/A
				NAG / Sand %	N/A
				SHS.FAME.neg / Clay %	N/A
				SHS.FAME.pos / Clay %	N/A
				SHS.FAME.fit / Clay %	N/A
				SHS.Enz.FAME.neg / Clay %	N/A
				SHS.Enz.FAME.pos / Clay %	N/A
				SHS.Enz.FAME.fit / Clay %	N/A

Table 4.4. Results [a] for 8 SH indicators under agricultural crop rotations, previous CRP conditions, and natural conditions in Ohio and Michigan.

Crop Rotation	n	Total Bacteria [c]	16:0 iso	18:0 10-ME	SHS (-) [FAME]	SHS (+) [FAME]	SHS (Fitted) [FAME]	SHS (Fitted) [Enz+FAME]	SHS (+) [Enz+FAME]
		nmol g^{-1} soil			-				
CS [b]	166	54.0 (±1.21) d [d]	4.76 (±0.113) d	2.86 (±0.064) e	0.383 (±0.007) d	0.374 (±0.007) d	0.379 (±0.007) d	0.404 (±0.006) d	0.409 (±0.006) d
CSW	80	63.1 (±1.43) c	5.66 (±0.160) c	3.38 (±0.084) d	0.461 (±0.011) c	0.458 (±0.011) c	0.460 (±0.011) c	0.466 (±0.010) c	0.469 (±0.009) c
CSW - Michigan	24	33.8 (±1.49) e	3.17 (±0.131) e	2.45 (±0.115) e	0.308 (±0.012) e	0.307 (±0.013) d	0.308 (±0.012) e	0.305 (±0.010) e	0.320 (±0.010) e
P	16	121 (±7.45) b	11.1 (±0.651) b	6.09 (±0.543) c	0.989 (±0.068) b	0.960 (±0.063) b	0.975 (±0.065) b	0.886 (±0.047) b	0.844 (±0.042) b
CRP	6	178 (±17.4) a	17.8 (±1.53) a	9.84 (±0.737) a	1.02 (±0.072) b	1.07 (±0.075) ab	1.04 (±0.073) b	1.03 (±0.059) a	1.01 (±0.052) a
VS	9	157 (±14) a	17.0 (±1.72) a	8.31 (±0.601) b	1.24 (±0.075) a	1.19 (±0.053) a	1.21 (±0.062) a	1.13 (±0.062) a	1.07 (±0.051) a

a ... Mean data observations and standard error values

b ... CS…Corn/Soy crop rotation; CSW…Corn/Soy/Wheat crop rotation; P… Prairies, CRP… Cons. Reserve Prog., VS… Virgin Soil

c ... Sum of 15:0, 17:0, a15:0, i15:0, i16:0, a17:0, i17:0, cy17:0, cy19:0, 16:1 w7c, 18:1 w7c

d ... Values within a column followed by the same lower-case letter are not significantly different at $P< 0.05$

Table 4.5. Results [a] for 8 SH indicators with or without cover cropping in Ohio.

Cover Crop usage	n	Total Bacteria [c]	16:0 iso	18:0 10-ME	SHS (-) [FAME]	SHS (+) [FAME]	SHS (Fitted) [FAME]	SHS (Fitted) [Enz+FAME]	SHS (+) [Enz+FAME]
		nmol g^{-1} soil			-				
CC (No) [b]	156	52.4 (±1.08) a [d]	4.69 (±0.112) a	2.82 (±0.060) a	0.376 (±0.007) a	0.368 (±0.007) a	0.372 (±0.007) a	0.395 (±0.006) a	0.402 (±0.005) a
CC (Yes)	90	64.8 (±1.58) b	5.67 (±0.157) b	3.39 (±0.091) b	0.464 (±0.011) b	0.460 (±0.011) b	0.462 (±0.011) b	0.474 (±0.009) b	0.475 (±0.008) b

a ... Mean data observations and standard error values

b ... CS…Corn/Soy crop; CSW…Corn/Soy/Wheat

c ... Sum of 15:0, 17:0, a15:0, i15:0, i16:0, a17:0, i17:0, cy17:0, cy19:0, 16:1 w7c, 18:1 w7c

d ... Values within a column followed by the same lower-case letter are not significantly different at P< 0.05

Table 4.6. Results [a] for 8 SH indicators with or without manure **i**n Ohio

Manure usage	n	Total Bacteria [c]	16:0 iso	18:0 10-ME	SHS (-) [FAME]	SHS (+) [FAME]	SHS (Fitted) [FAME]	SHS (Fitted) [Enz+FAME]	SHS (+) [Enz+FAME]
		nmol g^{-1} soil			-				
Man (No) [b]	152	53.3 (±1.19) a [d]	4.77 (±0.12) a	2.87 (±0.062) a	0.387 (±0.007) a	0.38 (±0.008) a	0.384 (±0.007) a	0.408 (±0.007) a	0.413 (±0.006) a
Man (Yes)	94	62.8 (±1.49) b	5.5 (±0.15) b	3.28 (±0.091) b	0.442 (±0.011) b	0.436 (±0.011) b	0.439 (±0.011) b	0.451 (±0.009) b	0.453 (±0.008) b

a ... Mean data observations and standard error values

b ... CS…Corn/Soy crop; CSW…Corn/Soy/Wheat

c ... Sum of 15:0, 17:0, a15:0, i15:0, i16:0, a17:0, i17:0, cy17:0, cy19:0, 16:1 w7c, 18:1 w7c

d ... Values within a column followed by the same lower-case letter are not significantly different at P< 0.05

Table 4.7. Results [a] for 8 SH indicators under various tillage systems in Ohio

Tillage system	n	Total Bacteria [c]	16:0 iso	18:0 10-ME	SHS (-) [FAME]	SHS (+) [FAME]	SHS (Fitted) [FAME]	SHS (Fitted) [Enz+FAME]	SHS (+) [Enz+FAME]
		_____ nmol g^{-1} soil _____			_____ - _____				
PT [b]	39	56 (±1.84) a [d]	4.94 (±0.195) a	2.98 (±0.107) a	0.411 (±0.013) a	0.408 (±0.013) a	0.409 (±0.013) a	0.421 (±0.011) a	0.426 (±0.01) a
MT	83	56.5 (±1.2) a	5.05 (±0.125) a	3.04 (±0.07) a	0.393 (±0.008) a	0.386 (±0.008) a	0.39 (±0.008) a	0.411 (±0.007) a	0.417 (±0.006) a
NT	124	57.6 (±1.67) a	5.09 (±0.16) a	3.03 (±0.089) a	0.417 (±0.011) a	0.41 (±0.011) a	0.414 (±0.011) a	0.434 (±0.01) a	0.438 (±0.009) a

a ... Mean data observations and standard error values

b ... CS…Corn/Soy crop; CSW…Corn/Soy/Wheat

c ... Sum of 15:0, 17:0, a15:0, i15:0, i16:0, a17:0, i17:0, cy17:0, cy19:0, 16:1 w7c, 18:1 w7c

d ... Values within a column followed by the same lower-case letter are not significantly different at P< 0.05

Table 4.8. Total Bacteria, 16:0 iso, 18:0 10-Me results [a] of farm fields with different management practices taken at a soil depth of 15 cm in Ohio.

Tillage practice	Crop Rotation	Cover Crop usage	Manure usage	n	Total Bacteria [d]		16:0 iso		18:0 10-ME	
					nmol g^{-1} soil					
PT [b]	CS [c]	No	No	6	43.8 (±1.73) CD [e]	b [f]	3.79 (±0.130) CD	a	2.46 (±0.070) BCD	a
	CSW	No	No	8	56.7 (±3.35) ABCD	ab	4.65 (±0.199) ABCD	a	2.87 (±0.161) ABCD	a
			Yes	9	57.3 (±4.26) ABCD	ab	5.46 (±0.584) ABCD	a	3.01 (±0.277) ABCD	a
		Yes	No	2	60.1 (±0.90) ABCD	ab	5.29 (±0.254) ABCD	a	2.96 (±0.005) ABCD	a
			Yes	14	59.4 (±3.22) ABC	a	5.20 (±0.305) ABCD	a	3.25 (±0.197) ABC	a
MT	CS	No	No	34	52.6 (±1.61) CD	c	4.74 (±0.135) BCD	b	2.99 (±0.080) BC	bc
			Yes	18	53.3 (±2.44) CD	c	4.54 (±0.292) BCD	b	2.65 (±0.174) CD	c
		Yes	No	6	51.7 (±4.45) BCD	bc	4.57 (±0.467) ABCD	b	2.70 (±0.257) BCD	bc
	CSW	No	No	10	66.7 (±2.48) ABC	a	6.65 (±0.291) A	a	3.66 (±0.162) AB	a
			Yes	13	63.6 (±2.72) ABC	ab	5.50 (±0.222) ABC	ab	3.32 (±0.150) ABC	ab
		Yes	Yes	2	67.3 (±1.08) ABCD	abc	5.45 (±0.162) ABCD	ab	3.58 (±0.027) ABCD	abc
NT	CS	No	No	52	44.8 (±2.02) D	b	4.01 (±0.211) D	b	2.41 (±0.104) D	b
			Yes	6	61.5 (±5.97) ABCD	ab	5.58 (±0.593) ABCD	ab	3.43 (±0.355) ABCD	ab
		Yes	No	20	64.3 (±2.80) ABC	a	5.59 (±0.271) ABC	a	3.20 (±0.175) ABC	a
			Yes	24	68.8 (±3.49) AB	a	5.94 (±0.311) AB	a	3.53 (±0.192) AB	a
	CSW	Yes	No	14	63.2 (±4.62) ABC	a	5.69 (±0.514) ABC	a	3.49 (±0.217) ABC	a
			Yes	8	77.0 (±4.04) A	a	6.82 (±0.617) A	a	4.10 (±0.364) A	a

a ... Data are means and standard error values

b ... PT…Plow Till; MT…Minimal-Till; NT…No-Till

c ... CS…Corn/Soy crop rotation; CSW…Corn/Soy/Wheat crop rotation

d ... Sum of 15:0, 17:0, a15:0, i15:0, i16:0, a17:0, i17:0, cy17:0, cy19:0, 16:1 w7c, 18:1 w7c

e ... Values followed by the same uppercase letters across all management systems are not significantly different at P< 0.05.

f ... Values within a column and tillage group followed by the same lower case letter are not significantly different at P<0.05.

Table 4.9. The Soil Health Scores [a] calculated using key EL-FAME variables. The classification of agricultural management systems is based on farm survey data. The total number of 246 observations represents soil samples (0-15 cm) that were collected from farm fields in Ohio.

Tillage practice	Crop Rotation	Cover Crop usage	Manure usage	n	SHS (-) [FAME]		SHS (+) [FAME]		SHS (Fitted) [FAME]	
PT [b]	CS [c]	No	No	6	0.335 (±0.012) CDE [d]	a [e]	0.339 (±0.015) CDE	a	0.337 (±0.013) CDE	a
	CSW	No	No	8	0.383 (±0.018) BCDE	a	0.378 (±0.019) BCDE	a	0.381 (±0.018) BCDE	a
			Yes	9	0.430 (±0.030) BCDE	a	0.425 (±0.029) BCDE	a	0.428 (±0.029) BCDE	a
		Yes	No	2	0.455 (±0.053) ABCDE	a	0.440 (±0.043) ABCDE	a	0.448 (±0.048) ABCDE	a
			Yes	14	0.440 (±0.024) BCD	a	0.439 (±0.023) BCD	a	0.440 (±0.023) BCD	a
MT	CS	No	No	34	0.374 (±0.010) DE	bc	0.368 (±0.010) DE	b	0.371 (±0.010) CDE	b
			Yes	18	0.355 (±0.018) DE	c	0.343 (±0.018) DE	b	0.349 (±0.018) DE	b
		Yes	No	6	0.383 (±0.020) BCDE	abc	0.375 (±0.019) BCDE	ab	0.379 (±0.019) BCDE	ab
	CSW	No	No	10	0.477 (±0.021) ABC	a	0.452 (±0.019) ABCD	a	0.465 (±0.020) ABC	a
			Yes	13	0.433 (±0.021) BCD	ab	0.439 (±0.02) BCD	a	0.436 (±0.021) BCD	a
		Yes	Yes	2	0.425 (±0.011) ABCDE	abc	0.445 (±0.011) ABCDE	ab	0.435 (±0.011) ABCDE	ab
NT	CS	No	No	52	0.343 (±0.012) E	c	0.331 (±0.012) E	c	0.337 (±0.012) E	c
			Yes	6	0.405 (±0.035) BCDE	bc	0.400 (±0.034) BCDE	bc	0.402 (±0.034) BCDE	bc
		Yes	No	20	0.422 (±0.017) BCD	b	0.416 (±0.019) BCD	b	0.419 (±0.018) BCD	b
			Yes	24	0.477 (±0.019) B	b	0.468 (±0.019) ABC	b	0.473 (±0.019) AB	b
	CSW	Yes	No	14	0.490 (±0.028) AB	ab	0.502 (±0.030) AB	ab	0.496 (±0.029) AB	ab
			Yes	8	0.594 (±0.043) A	a	0.579 (±0.044) A	a	0.587 (±0.043) A	a

a ... Data are means and standard error values

b ... PT…Plow Till; MT…Minimal-Till; NT…No-Till

c ... CS…Corn/Soy crop rotation; CSW…Corn/Soy/Wheat crop rotation

e ... Values followed by the same uppercase letters across all management systems are not significantly different at P< 0.05.

f ... Values within a column and tillage group followed by the same lower case letter are not significantly different at P<0.05.

Table 4.10. The Soil Health Scores [a] calculated using key Enzyme and EL-FAME variables. The classification of agricultural management systems is based on farm survey information. The total number of 246 observations represents soil samples (0-15 cm) that were collected at farm fields throughout Ohio.

Tillage practice	Crop Rotation	Cover Crop usage	Manure usage	n	SHS (Fitted) [Enz+FAME]		SHS (+) [Enz+FAME]	
PT [b]	CS [c]	No	No	6	0.355 (±0.009) CD [d]	b [e]	0.366 (±0.010) CD	b
	CSW	No	No	8	0.426 (±0.021) BCD	ab	0.426 (±0.018) BCD	ab
			Yes	9	0.415 (±0.026) BCD	ab	0.423 (±0.023) BCD	ab
		Yes	No	2	0.444 (±0.022) ABCD	ab	0.446 (±0.014) ABCD	ab
			Yes	14	0.446 (±0.018) BC	a	0.451 (±0.017) BC	a
MT	CS	No	No	34	0.394 (±0.009) CD	bc	0.403 (±0.008) CD	bc
			Yes	18	0.382 (±0.016) CD	c	0.388 (±0.014) CD	c
		Yes	No	6	0.414 (±0.028) BCD	abc	0.417 (±0.024) BCD	abc
	CSW	No	No	10	0.457 (±0.016) ABC	a	0.453 (±0.014) BCD	ab
			Yes	13	0.450 (±0.020) BC	ab	0.456 (±0.018) BC	a
		Yes	Yes	2	0.456 (±0.003) ABCD	abc	0.461 (±0.004) ABCD	abc
NT	CS	No	No	52	0.367 (±0.012) D	c	0.375 (±0.011) D	c
			Yes	6	0.430 (±0.026) BCD	bc	0.437 (±0.024) BCD	bc
		Yes	No	20	0.447 (±0.015) BC	b	0.449 (±0.014) BC	b
			Yes	24	0.482 (±0.018) AB	ab	0.478 (±0.016) AB	ab
	CSW	Yes	No	14	0.501 (±0.027) AB	ab	0.510 (±0.026) AB	ab
			Yes	8	0.578 (±0.033) A	a	0.569 (±0.030) A	a

a ... Data are means and standard error values

b ... PT…Plow Till; MT…Minimal-Till; NT…No-Till

c ... CS…Corn/Soy crop rotation; CSW…Corn/Soy/Wheat crop rotation

e ... Values followed by the same uppercase letters across all management systems are not significantly different at P< 0.05.

f ... Values within a column and tillage group followed by the same lower case letter are not significantly different at P<0.05.

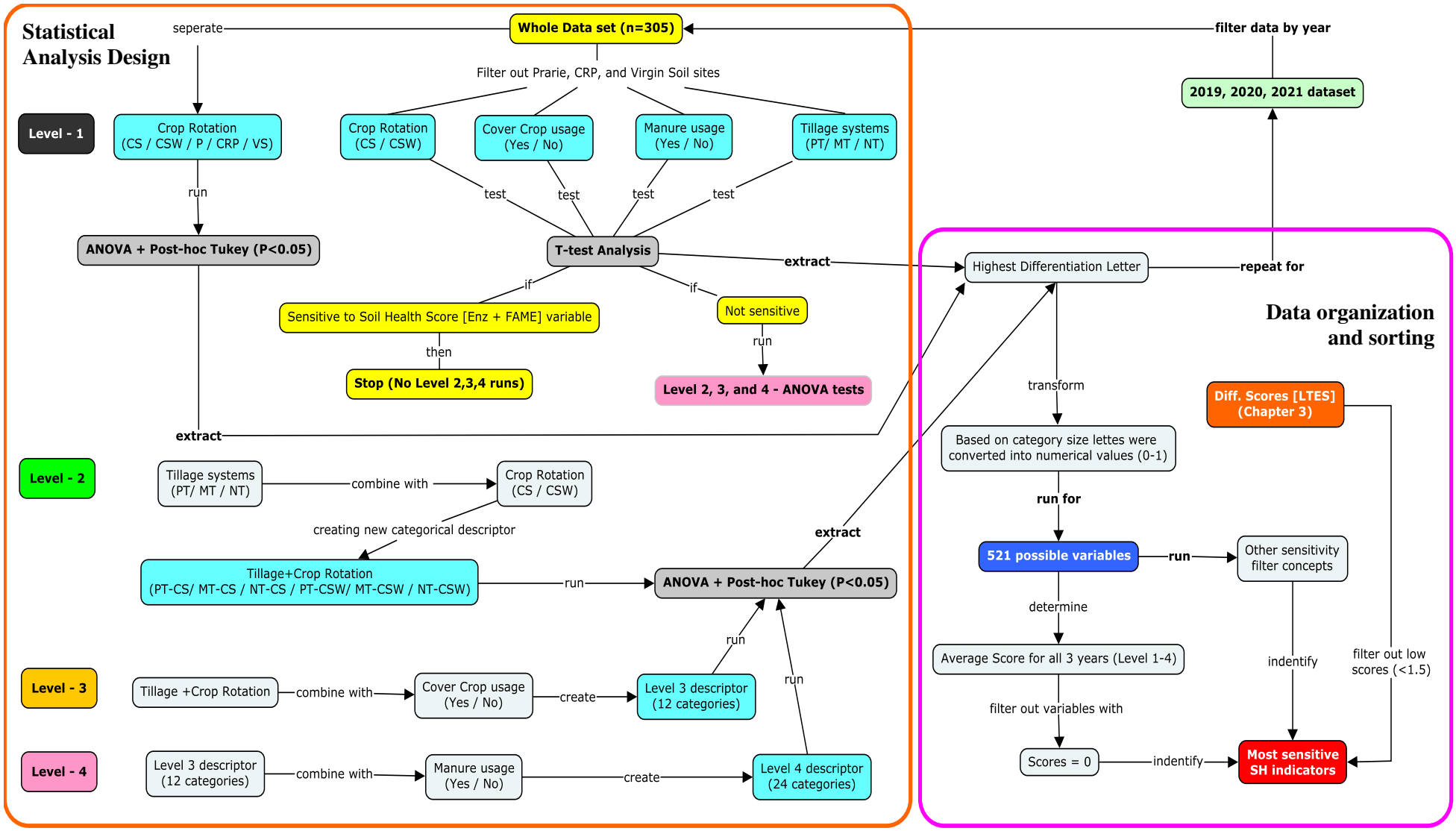


Figure 4.1. Flow chart of the multi-level sensitivity test and the calculation of differentiation power scores for 521 variables used to identify the most sensitive SH indicators. Light blue describes the 8 categorical classifications that were used.

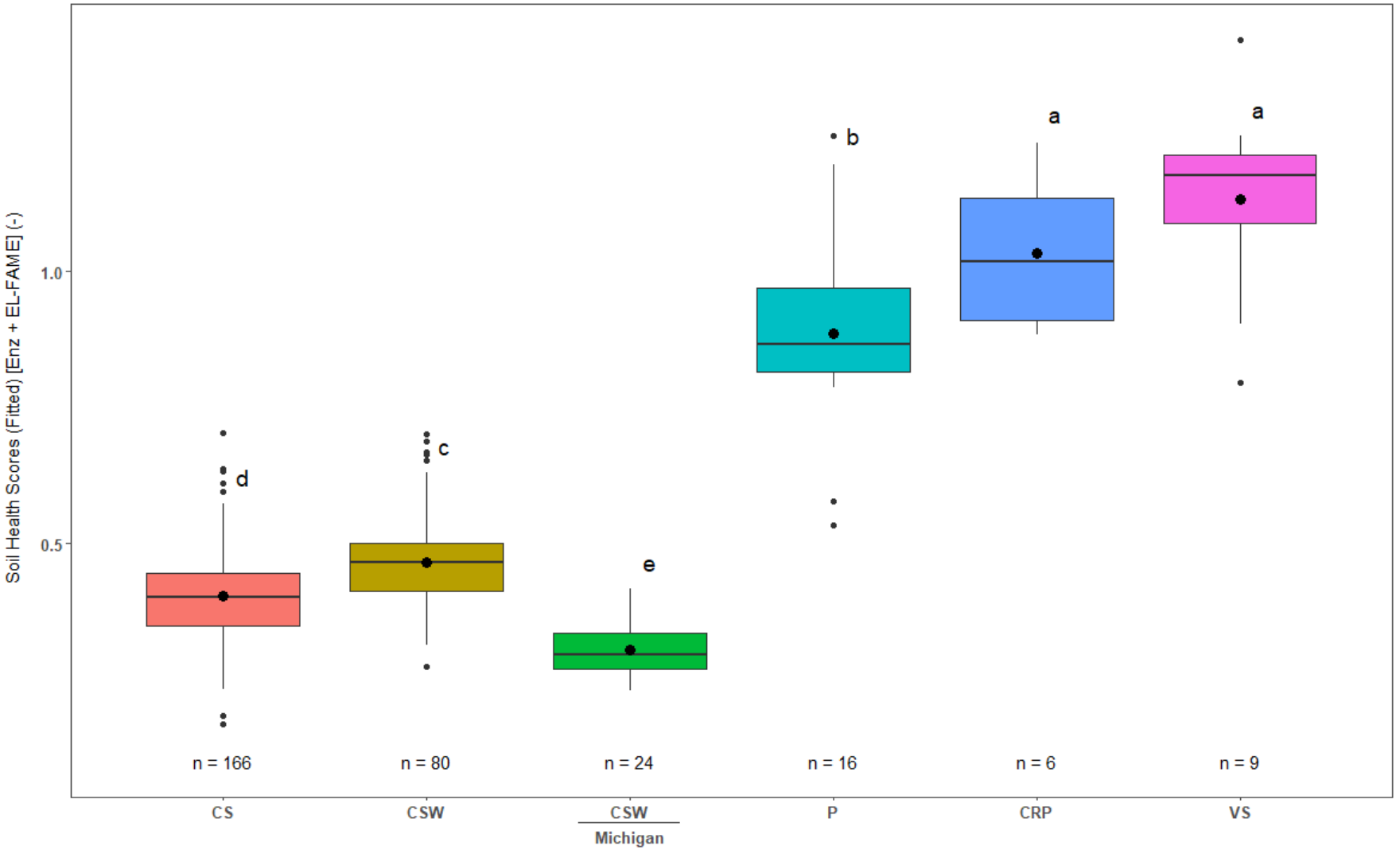


Figure 4.2. Effect of crop rotations, prairie, Conservation Reserve Program, and virgin Soil sites on Soil Health Scores (Fitted) [Enz + EL-FAME] variable)
CS... corn/soy, CSW...corn/soy/wheat, CSW-Michigan ... LTES in Michigan,
P... prairies, CRP... Conservation Reservation Program, VS... virgin soil.
Bars with the same lower-case letter are not significantly different at P<0.05

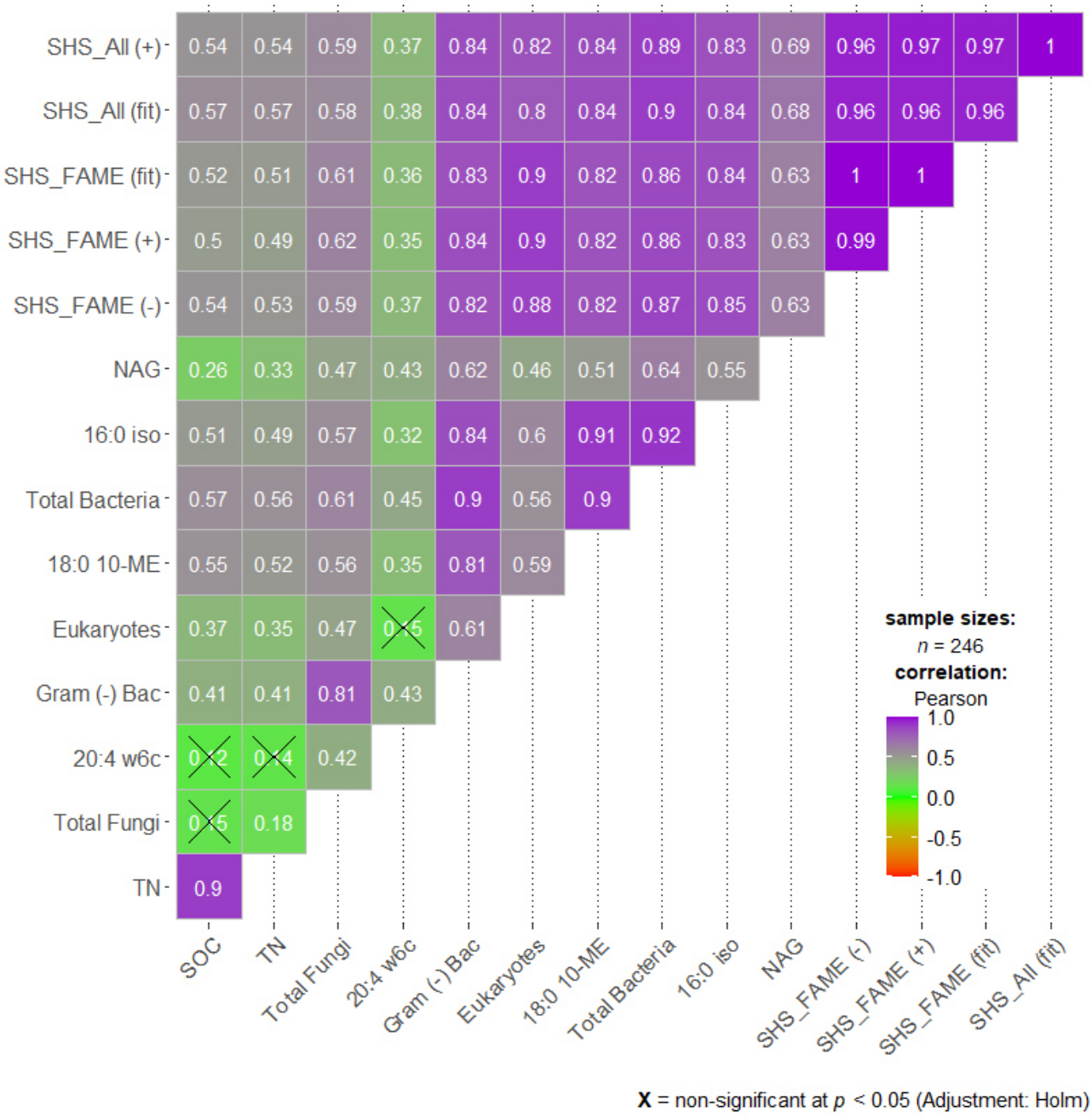


Figure 4.3. Correlation Matrix of 13 SH indicators identified in the multi-level sensitivity test that were used for develop differentiation scores (Chapter 3). Michigan LTES data excluded.

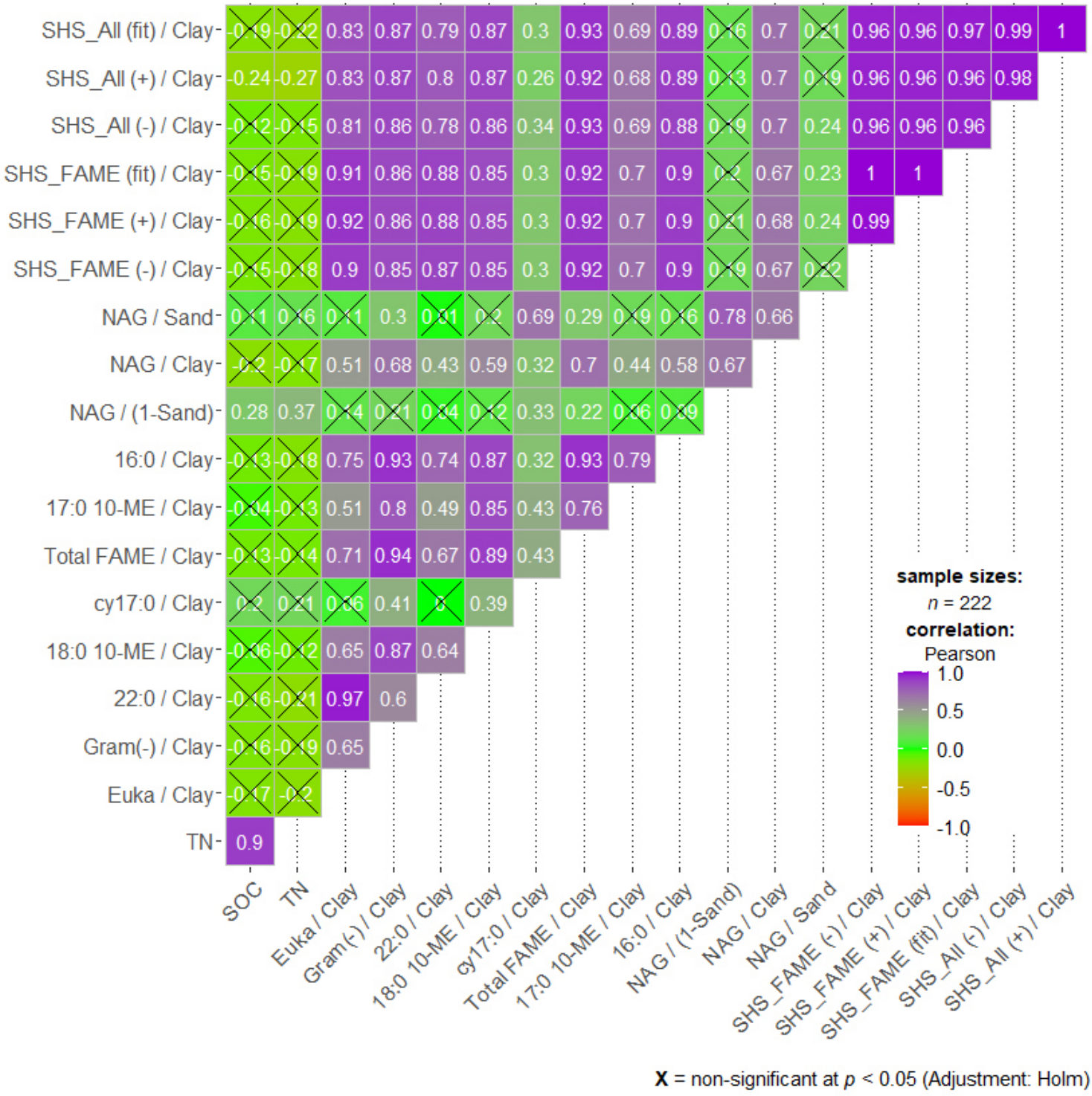


Figure 4.4. Correlation Matrix of the remaining 17 potential SH indicators identified in the multi-level sensitivity test. All LTES observations were excluded.

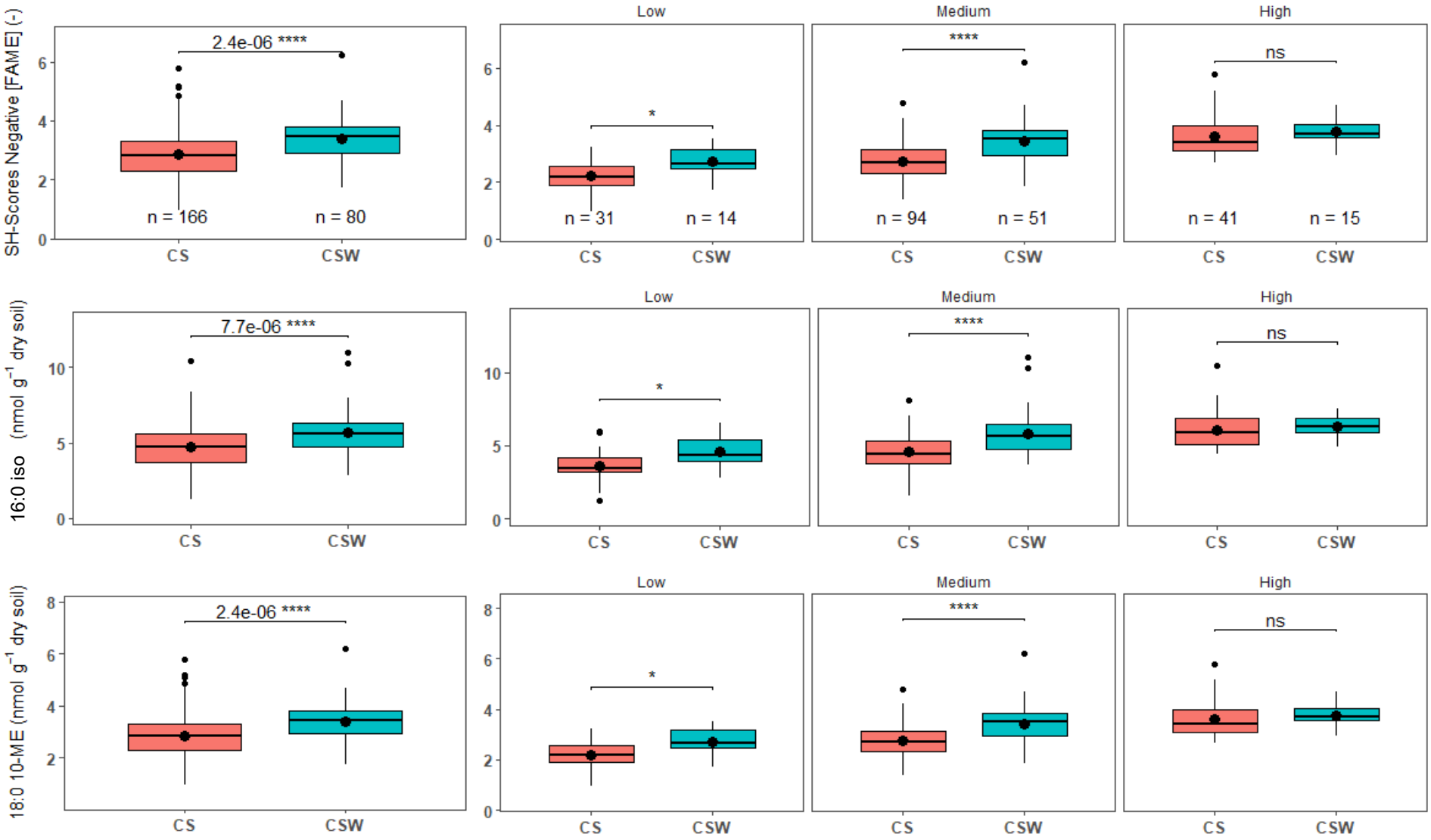
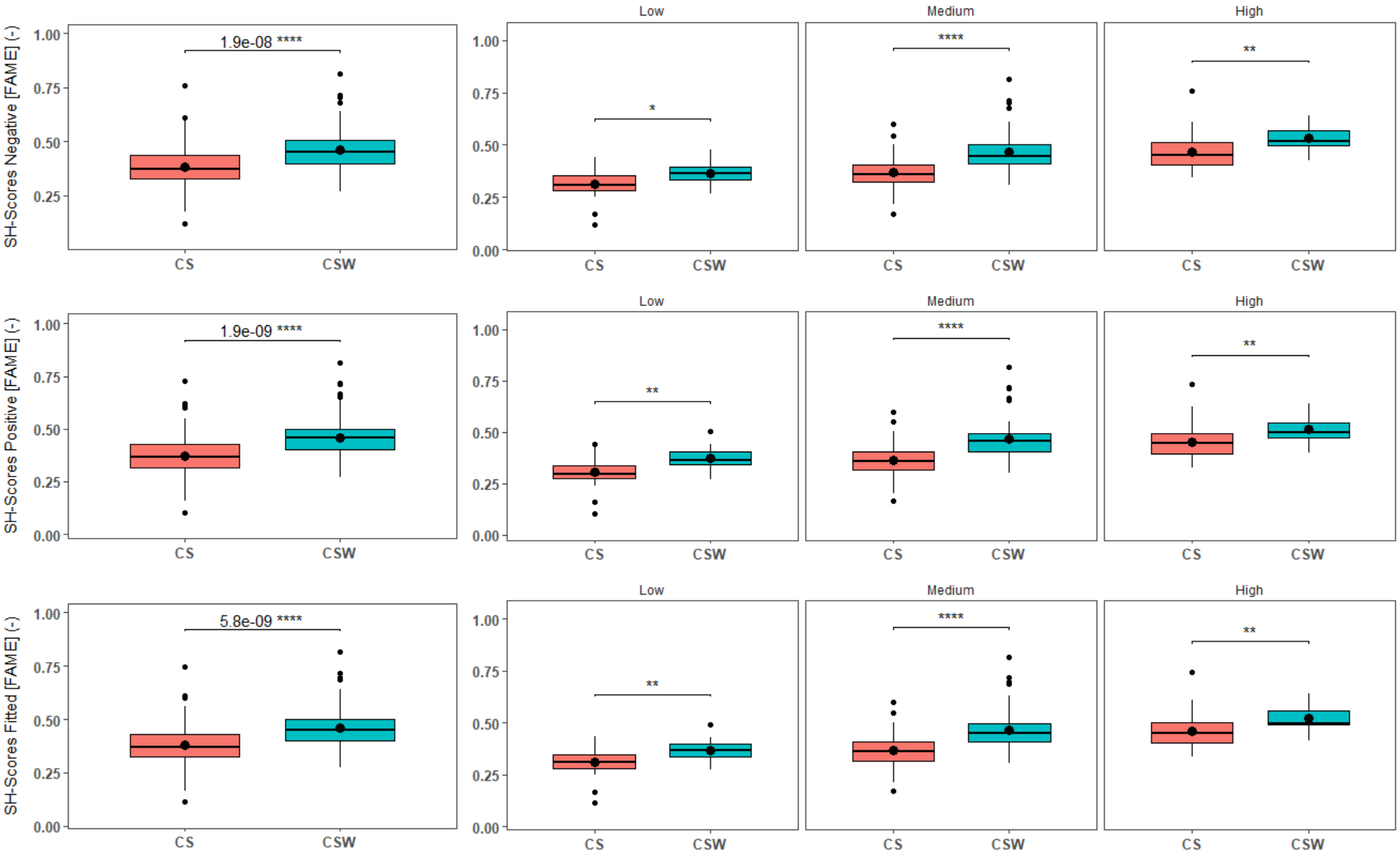
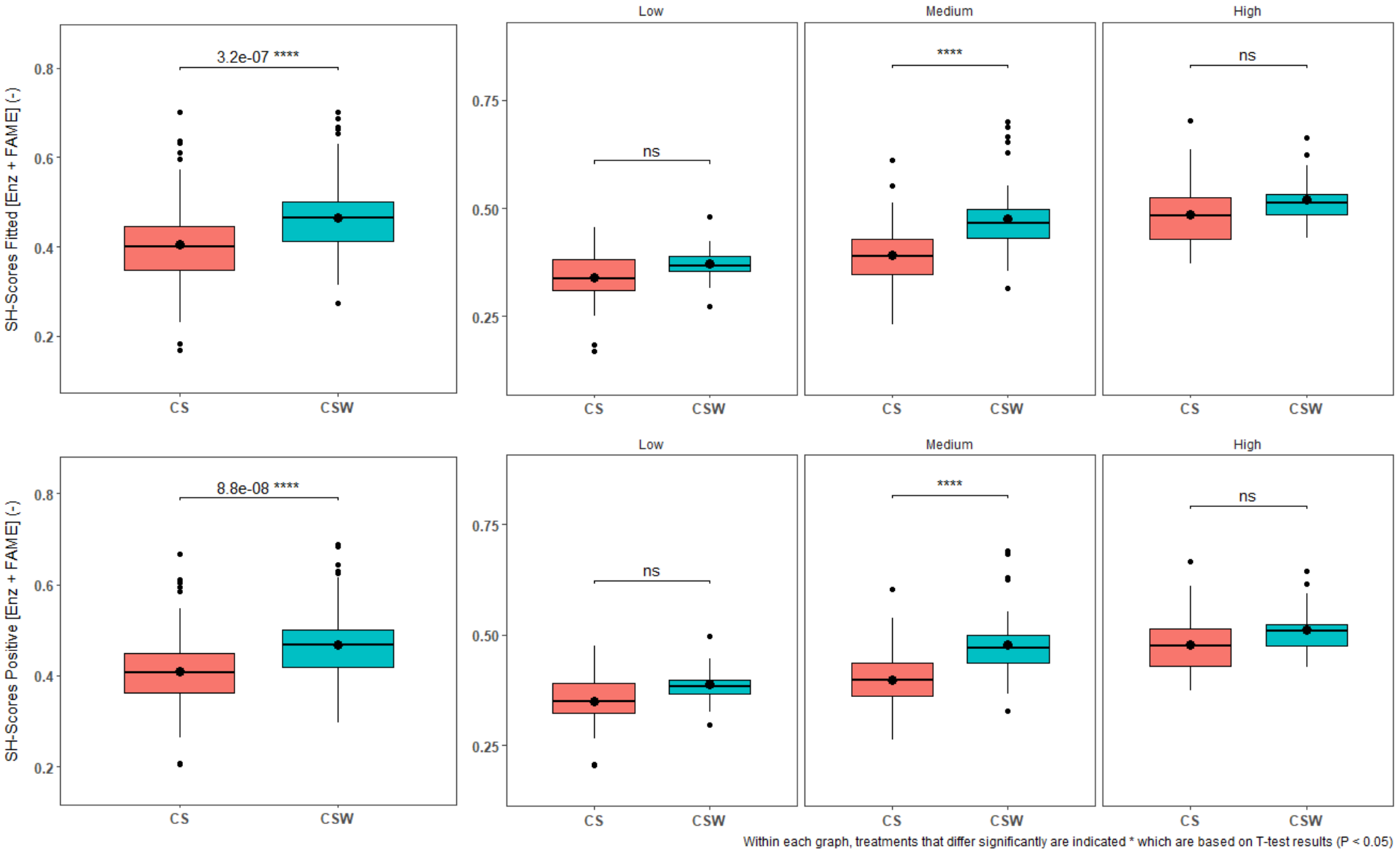


Figure 4.5. Eight SH indicator t-test results under corn-soy and corn-soy-wheat crop rotation conditions. The first graph for each variable represents the total number of observations. The next three graphs separate these observations into three SOC ranges [Low (<10.88g/kg), Medium (10.88g/kg ≥ 16.20g/kg), and High (>16.20g/kg)]. Michigan LTES observations were excluded (outlier).
Horizontal brackets over plots indicate statistical significance; *p < 0.05, **p < 0.01, and ***p < 0.001. ns=not significant.

continues


SH-Scores Negative [FAME] (-)
SH-Scores Positive [FAME] (-)
SH-Scores Fitted [FAME] (-)
1.00
0.75
0.50
0.25
0.00
1.9e-08 ****
1.9e-09 ****
5.8e-09 ****
Low
Medium
High
*
**
****
CS
CSW

continues



Within each graph, treatments that differ significantly are indicated * which are based on T-test results (P < 0.05)

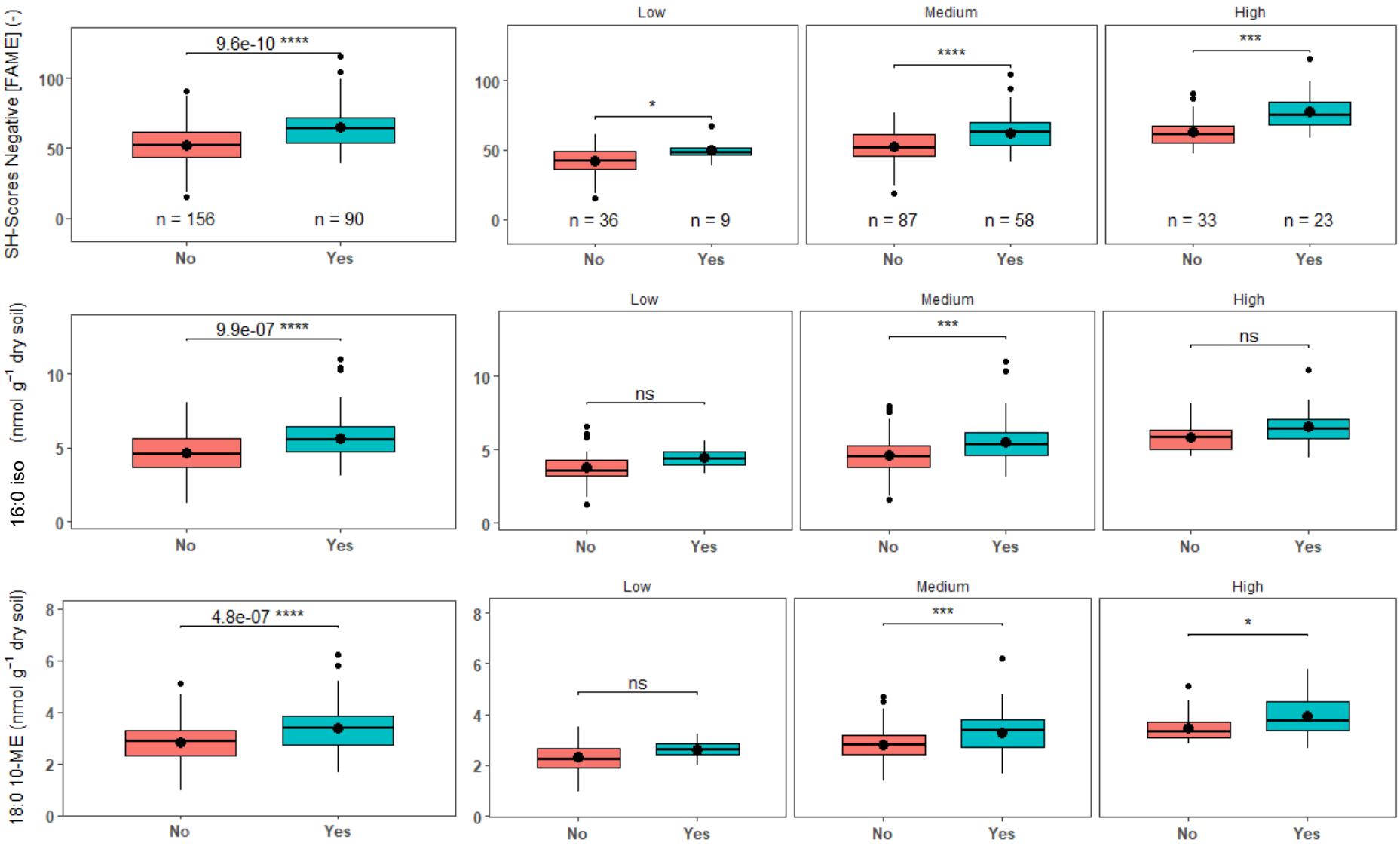
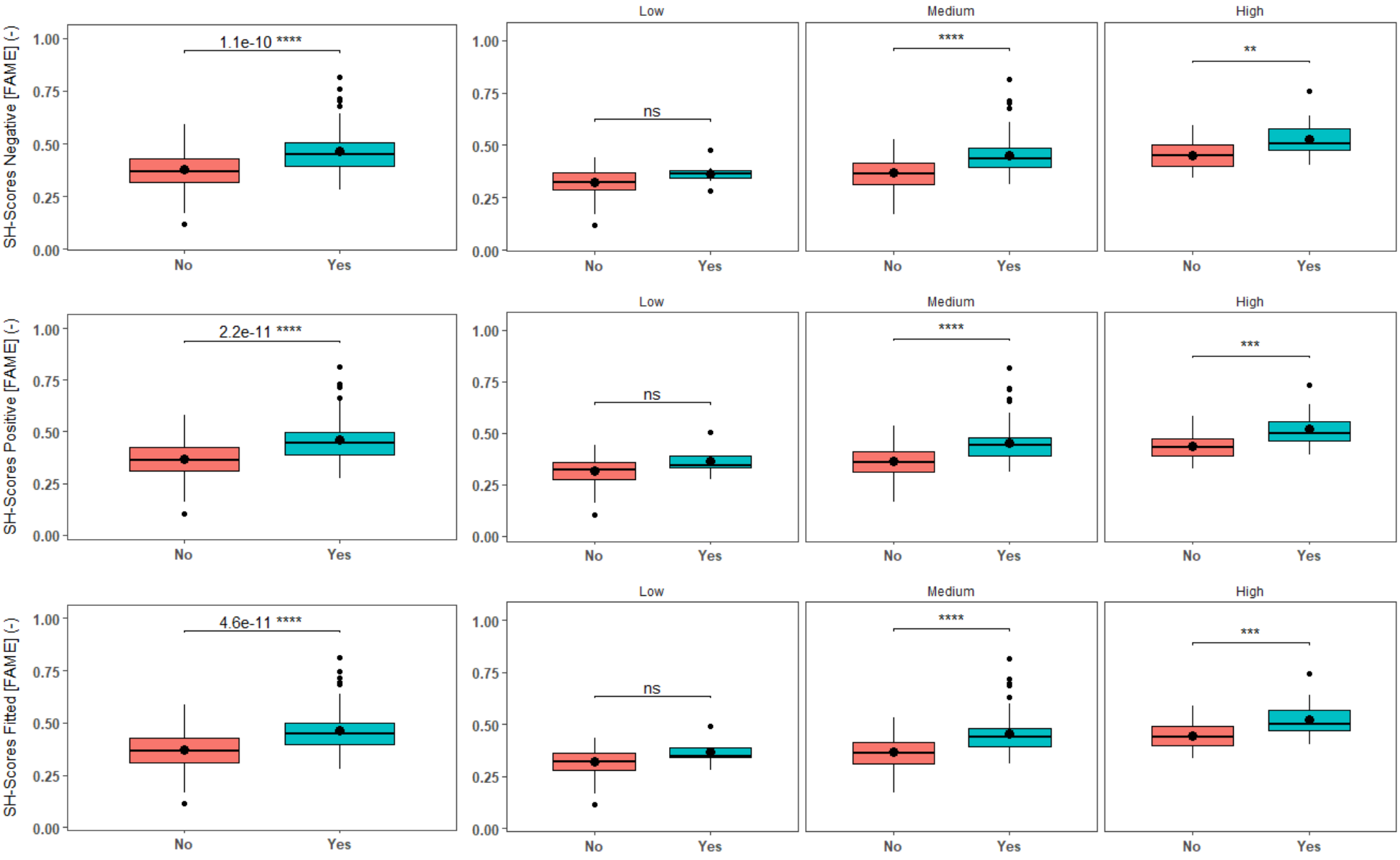
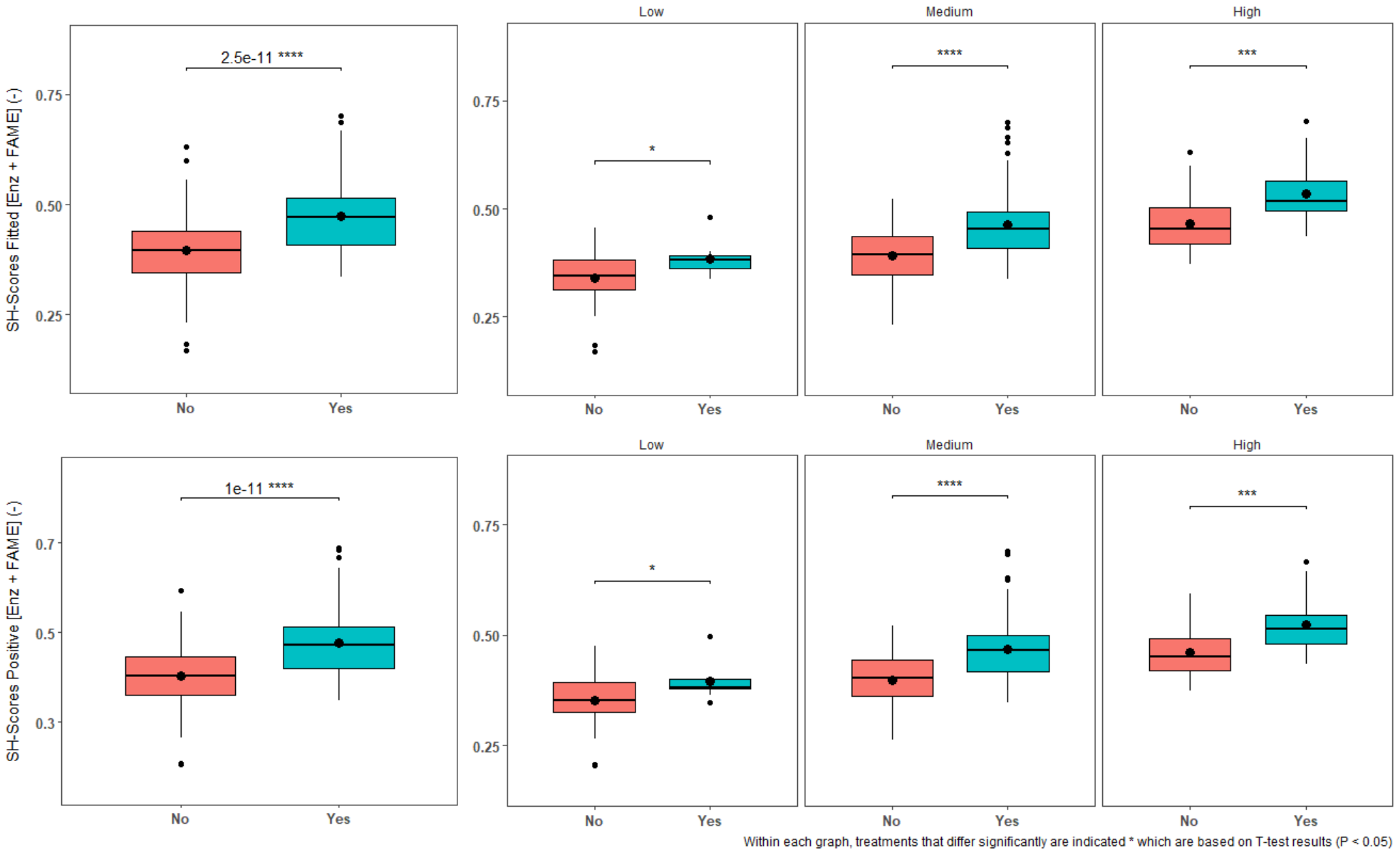


Figure 4.6. Eight SH indicator t-test results under cover cropping conditions. The first graph for each variable represents the total number of observations. The next three graphs separate these observations into three SOC ranges [Low (<10.88g/kg), Medium (10.88g/kg ≥ 16.20g/kg), and High (>16.20g/kg)]. Michigan LTES observations were excluded (outlier).
Horizontal brackets over plots indicate statistical significance; *p < 0.05, **p < 0.01, and ***p < 0.001. ns=not significant. **continues**


SH-Scores Negative [FAME] (-)
SH-Scores Positive [FAME] (-)
SH-Scores Fitted [FAME] (-)
1.1e-10 ****
2.2e-11 ****
4.6e-11 ****
Low
Medium
High
ns
****
**
***
No
Yes
0.00
0.25
0.50
0.75
1.00



Within each graph, treatments that differ significantly are indicated * which are based on T-test results (P < 0.05)

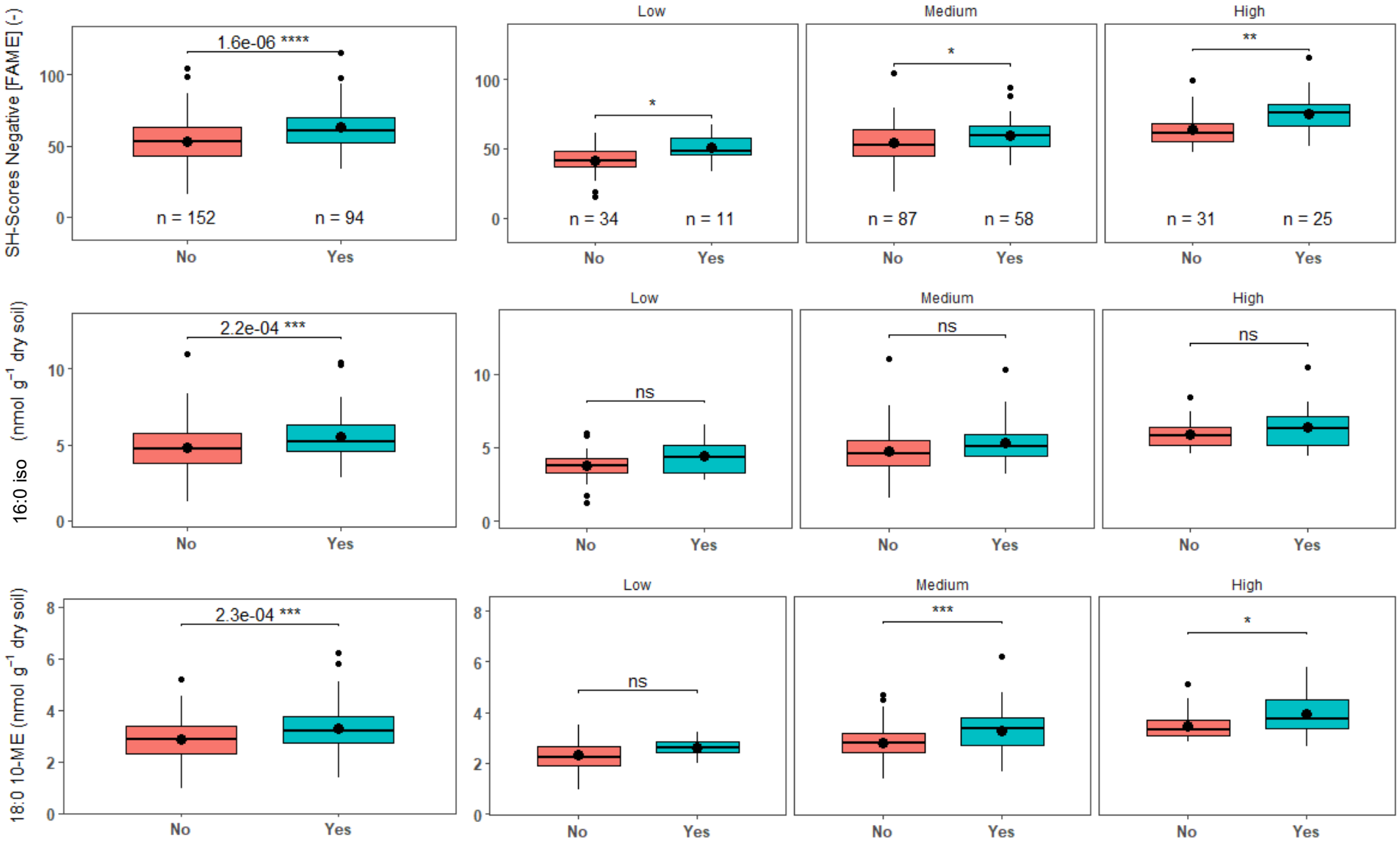
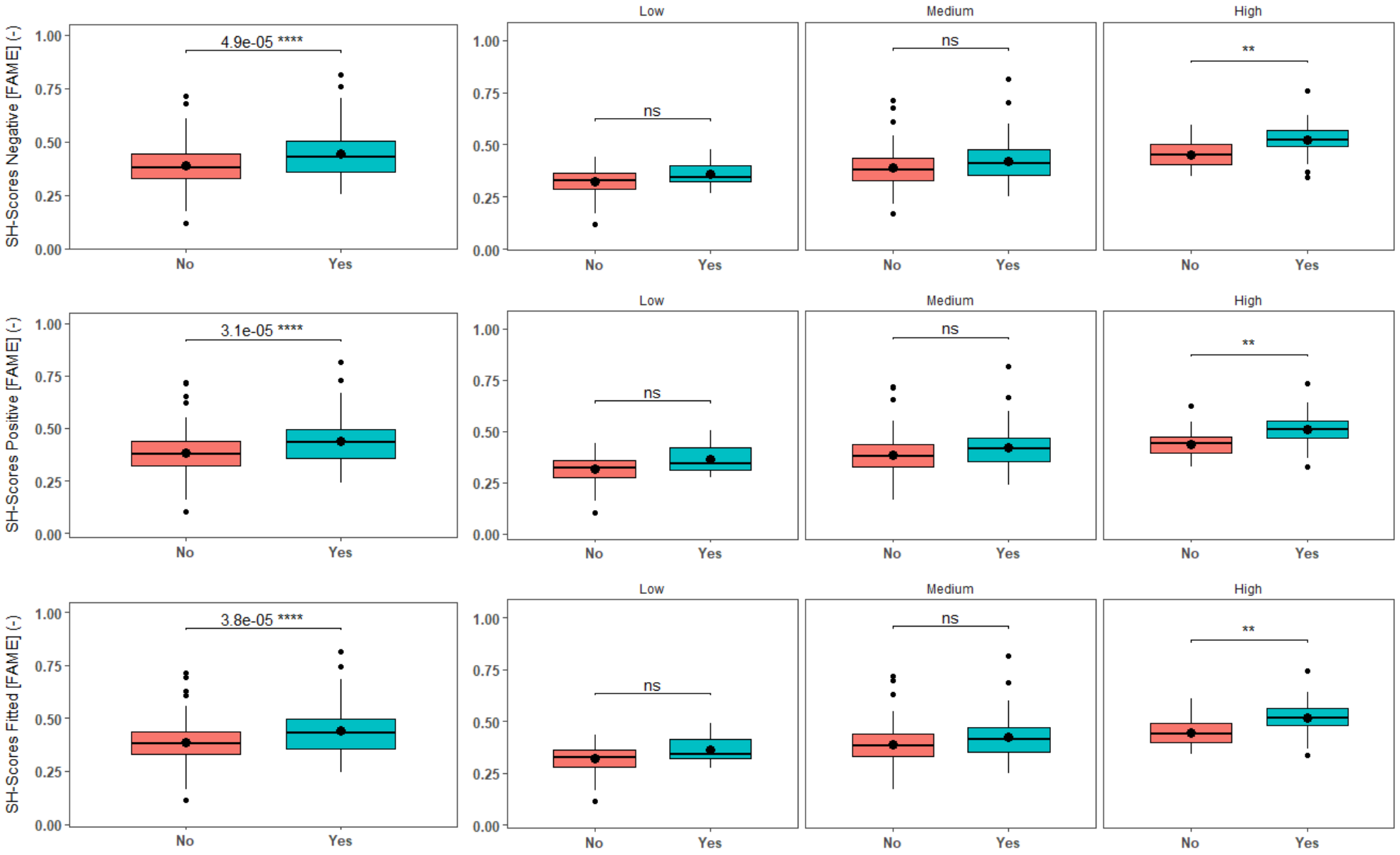
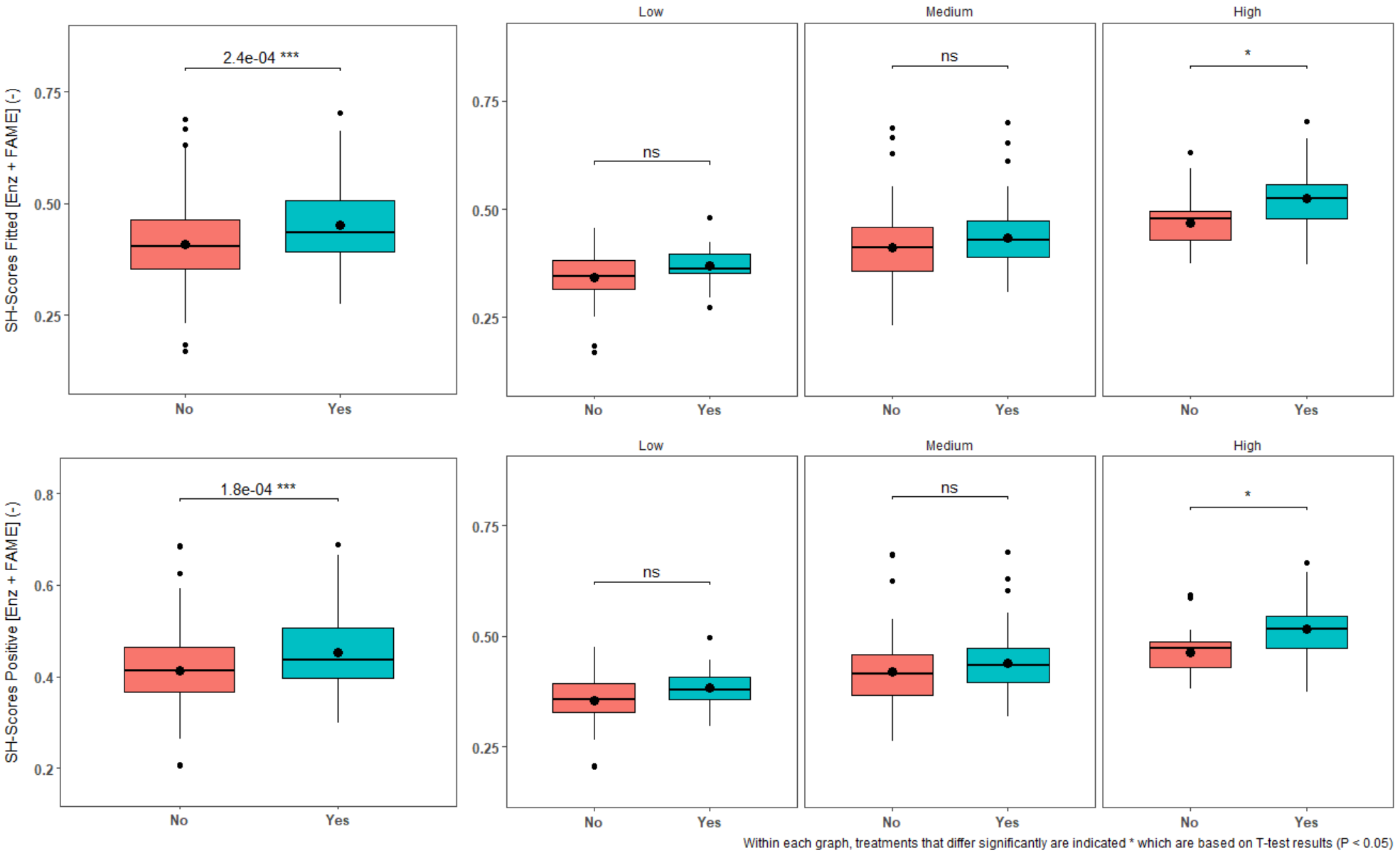


Figure 4.7. Eight SH indicator t-test results under manure conditions. The first graph for each variable represents the total number of observations. The next three graphs separate these observations into three SOC ranges [Low (<10.88g/kg), Medium (10.88g/kg ≥ 16.20g/kg), and High (>16.20g/kg)]. Michigan LTES observations were excluded (outlier).
Horizontal brackets over plots indicate statistical significance; *p < 0.05, **p < 0.01, and ***p < 0.001. ns=not significant. **continues**


SH-Scores Negative [FAME] (-)
SH-Scores Positive [FAME] (-)
SH-Scores Fitted [FAME] (-)
1.00
0.75
0.50
0.25
0.00
No
Yes
4.9e-05 ****
3.1e-05 ****
3.8e-05 ****
Low
Medium
High
ns
**

continues


SH-Scores Fitted [Enz + FAME] (-)
SH-Scores Positive [Enz + FAME] (-)
2.4e-04 ***
1.8e-04 ***
Low
Medium
High
ns
*
No
Yes
Within each graph, treatments that differ significantly are indicated * which are based on T-test results (P < 0.05)

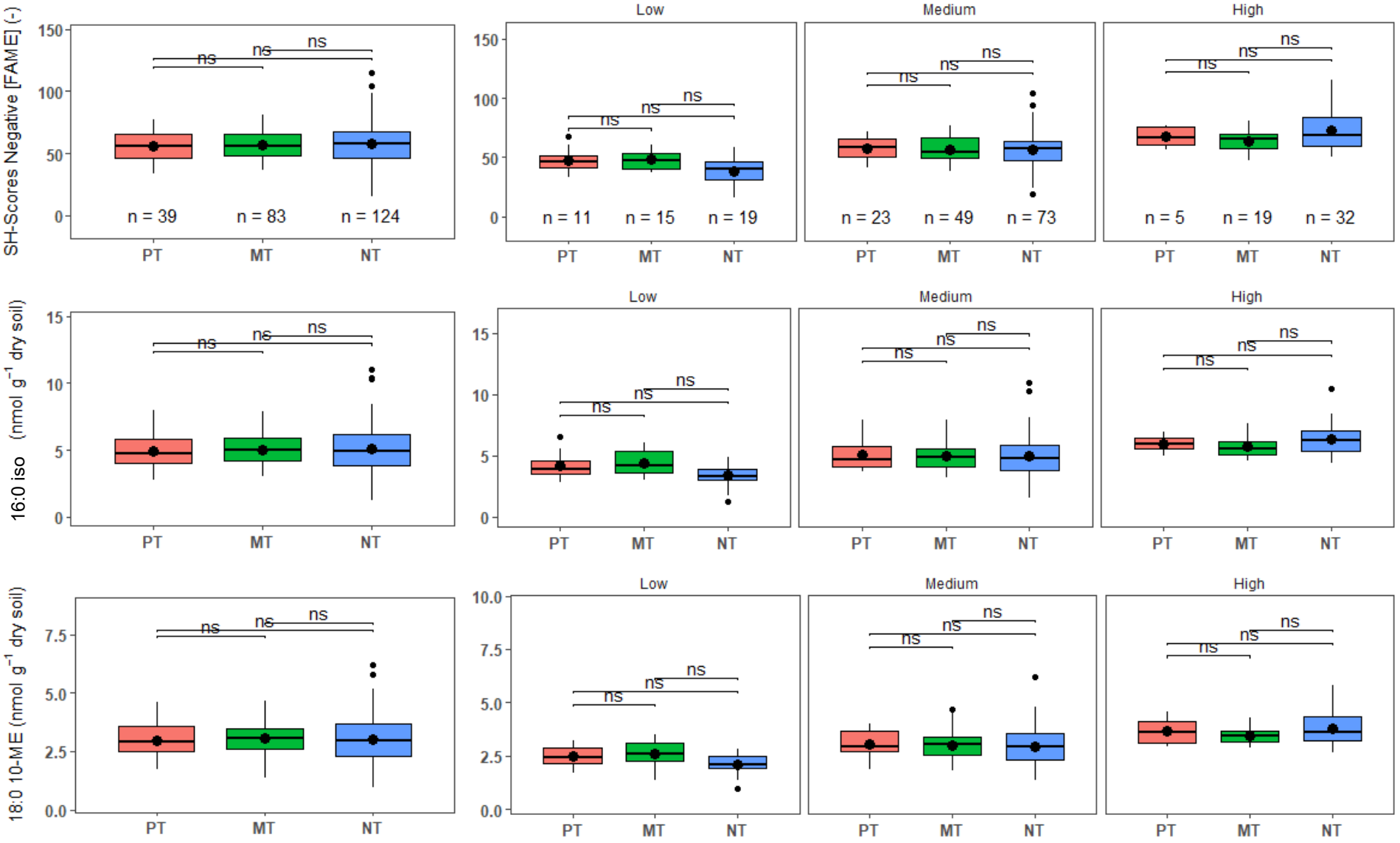
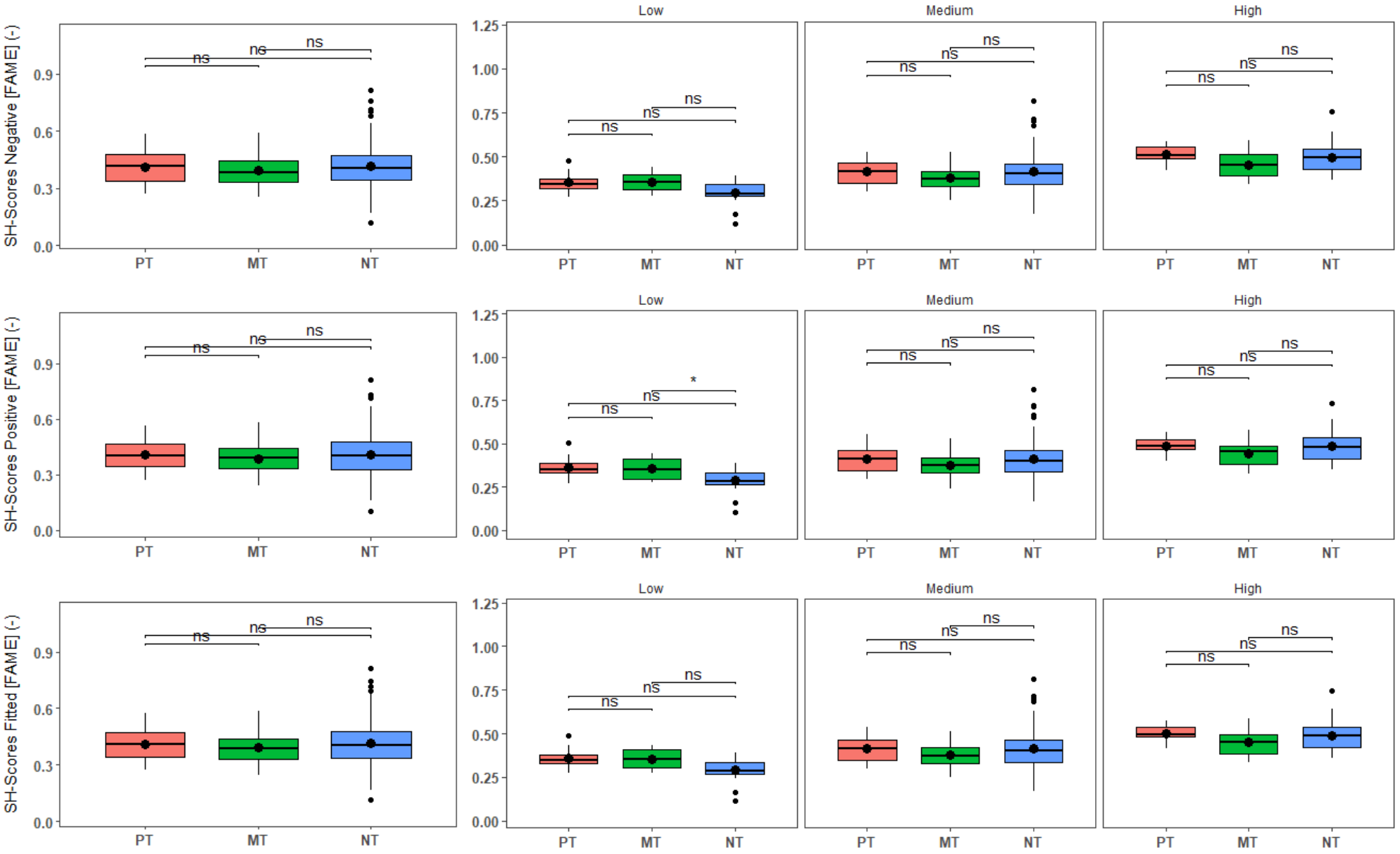
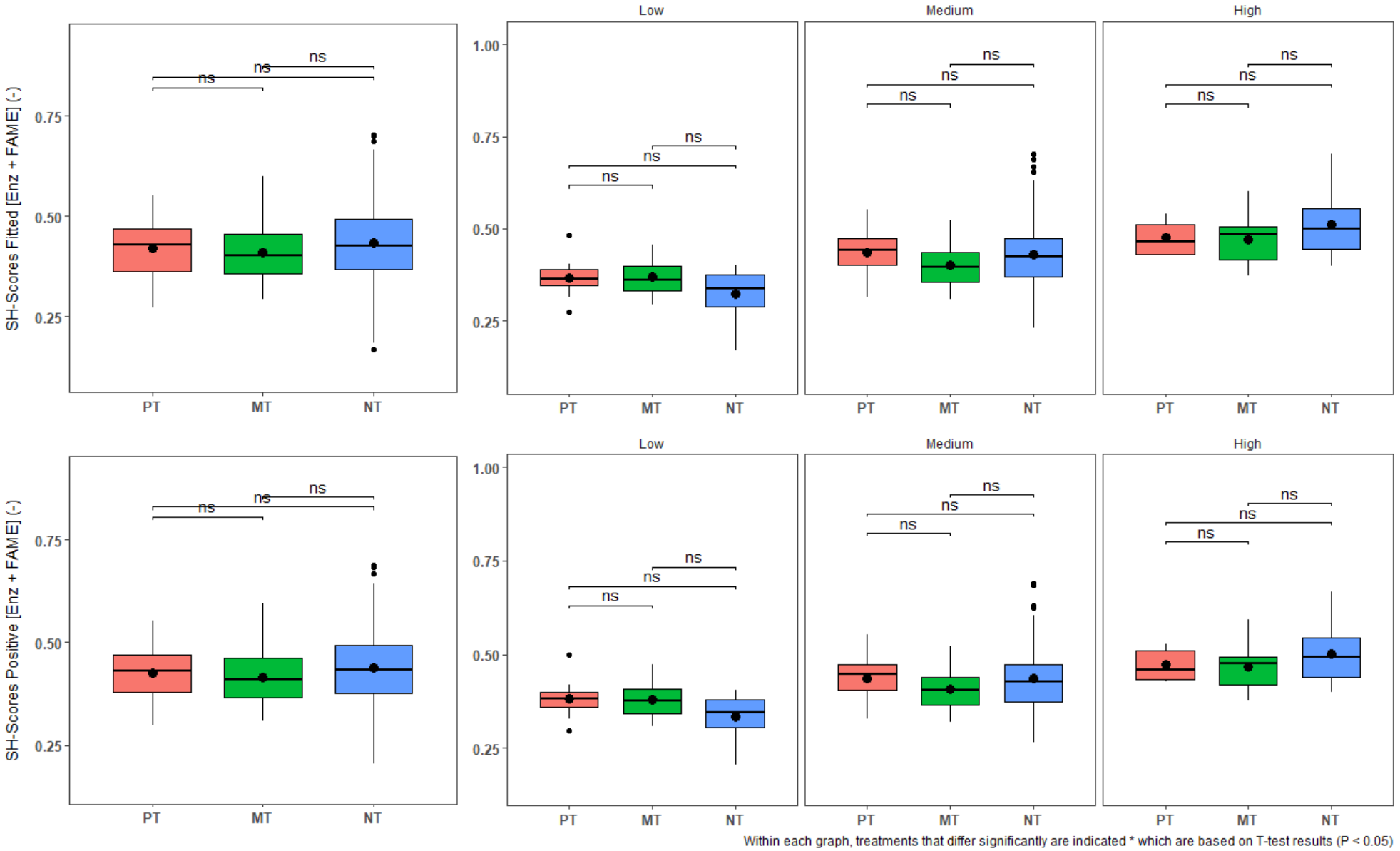


Figure 4.8. Eight SH indicator t-test results under various tillage systems. The first graph for each variable represents the total number of observations. The next three graphs separate these observations into three SOC ranges [Low (<10.88g/kg), Medium (10.88g/kg ≥ 16.20g/kg), and High (>16.20g/kg)]. Michigan LTES observations were excluded (outlier).
Horizontal brackets over plots indicate statistical significance; *p < 0.05, **p < 0.01, and ***p < 0.001. ns=not significant. **continues**


Low
Medium
High
SH-Scores Negative [FAME] (-)
SH-Scores Positive [FAME] (-)
SH-Scores Fitted [FAME] (-)
PT
MT
NT
ns
*

Continues


Low
Medium
High
SH-Scores Fitted [Enz + FAME] (-)
SH-Scores Positive [Enz + FAME] (-)
0.25
0.50
0.75
1.00
ns
PT
MT
NT
Within each graph, treatments that differ significantly are indicated * which are based on T-test results (P < 0.05)

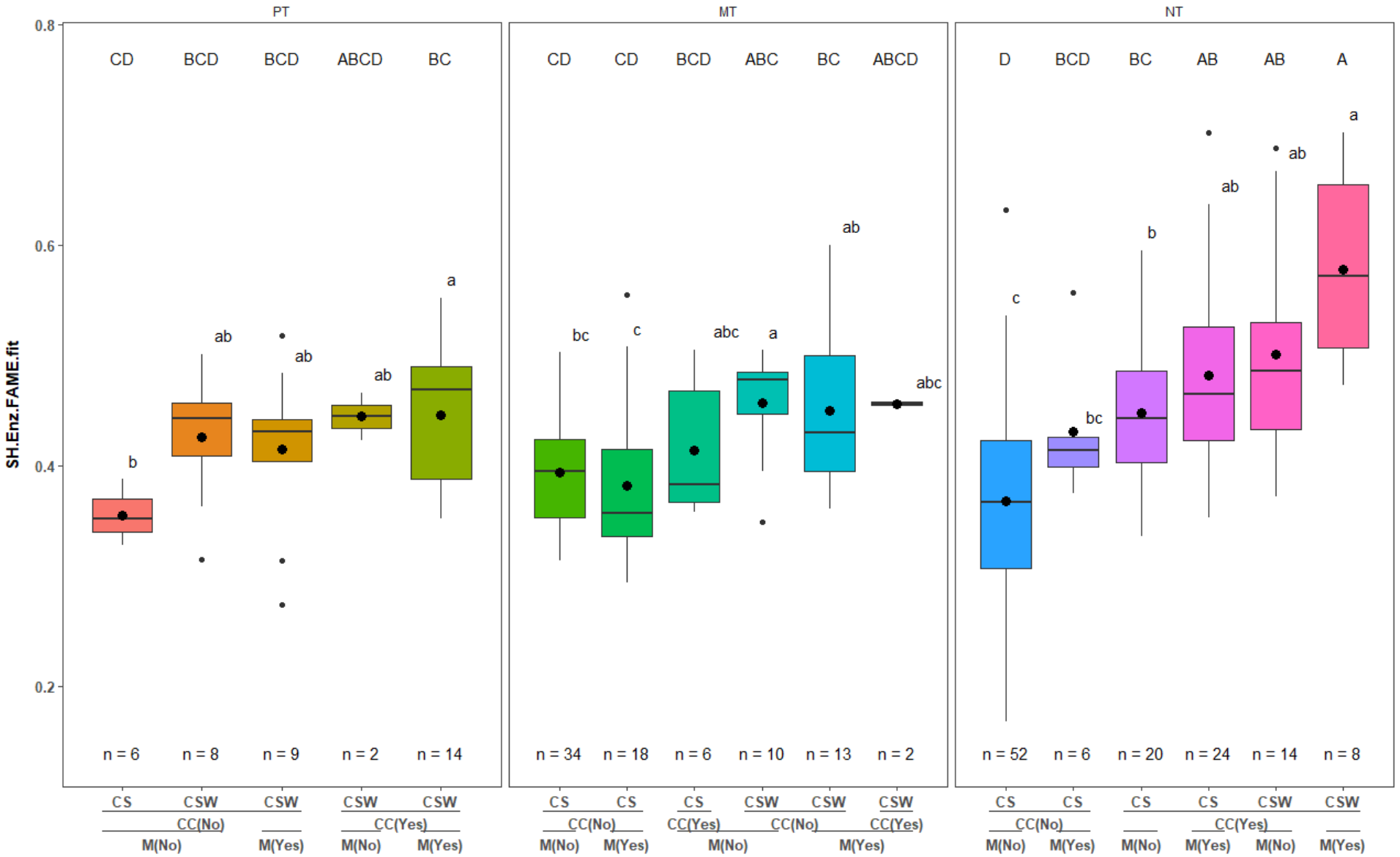


Figure 4.9. Shifts in SH Scores [Enzyme + EL-FAME] due to agricultural management systems. The graphs are separated by tillage systems (PT, MT, NT) and observations are categorized by crop rotation (CS, CSW) and further separated by cover crop usage (CC(No), CC(Yes)) and by manure usage (M(No), M(Yes)). Bars across all graphs followed the same uppercase letters are not significantly different at P<0.05. Bars within a graph with the same lowercase letter are not significantly different at p<0.05.

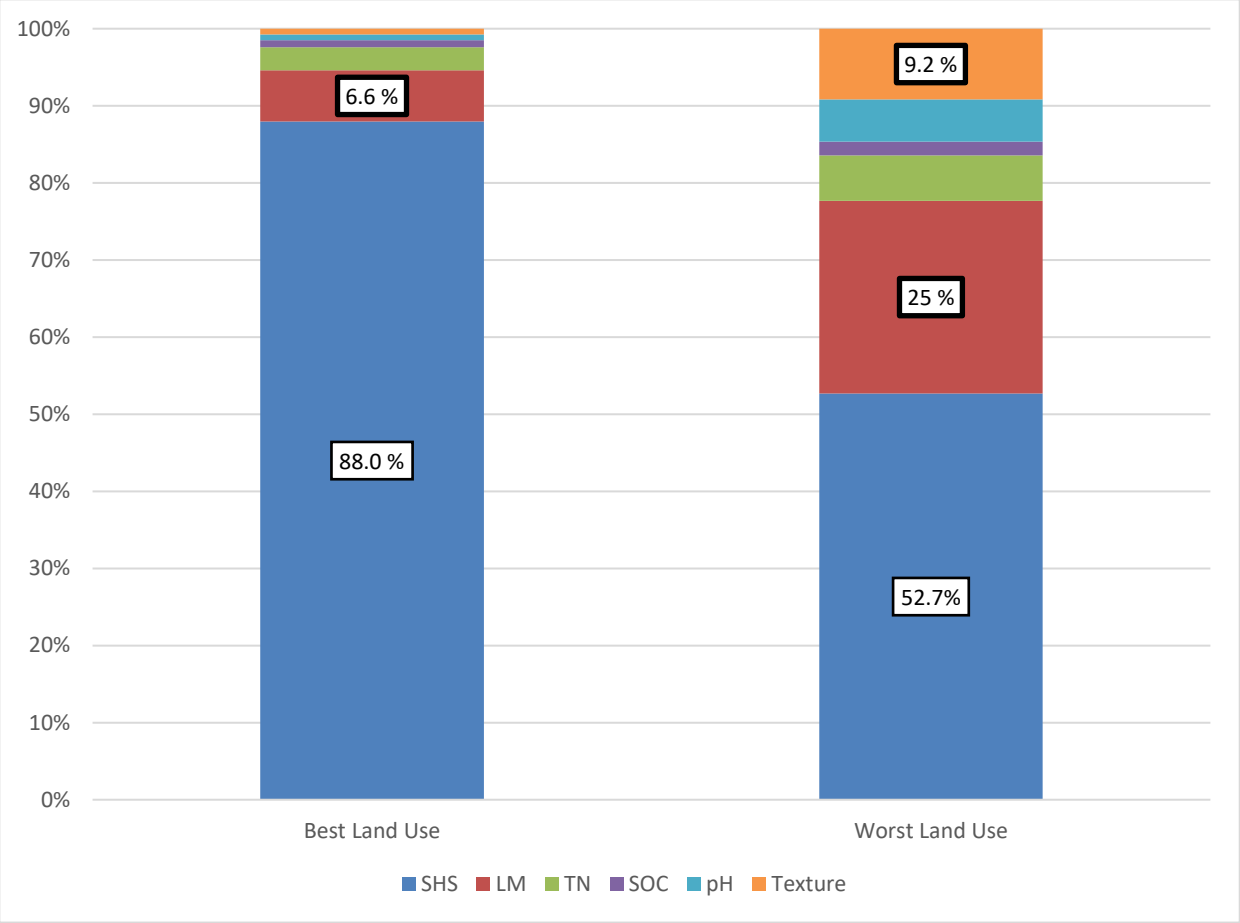


Figure 4.10. Percentage distribution of the absolute coefficient values based on the best and worst land management and soil health scenario.

SHS... Soil Health Score
LM... Land Management Practices
TN... Total Nitrogen
SOC... Soil organic Carbon
pH characteristics
Texture characteristic

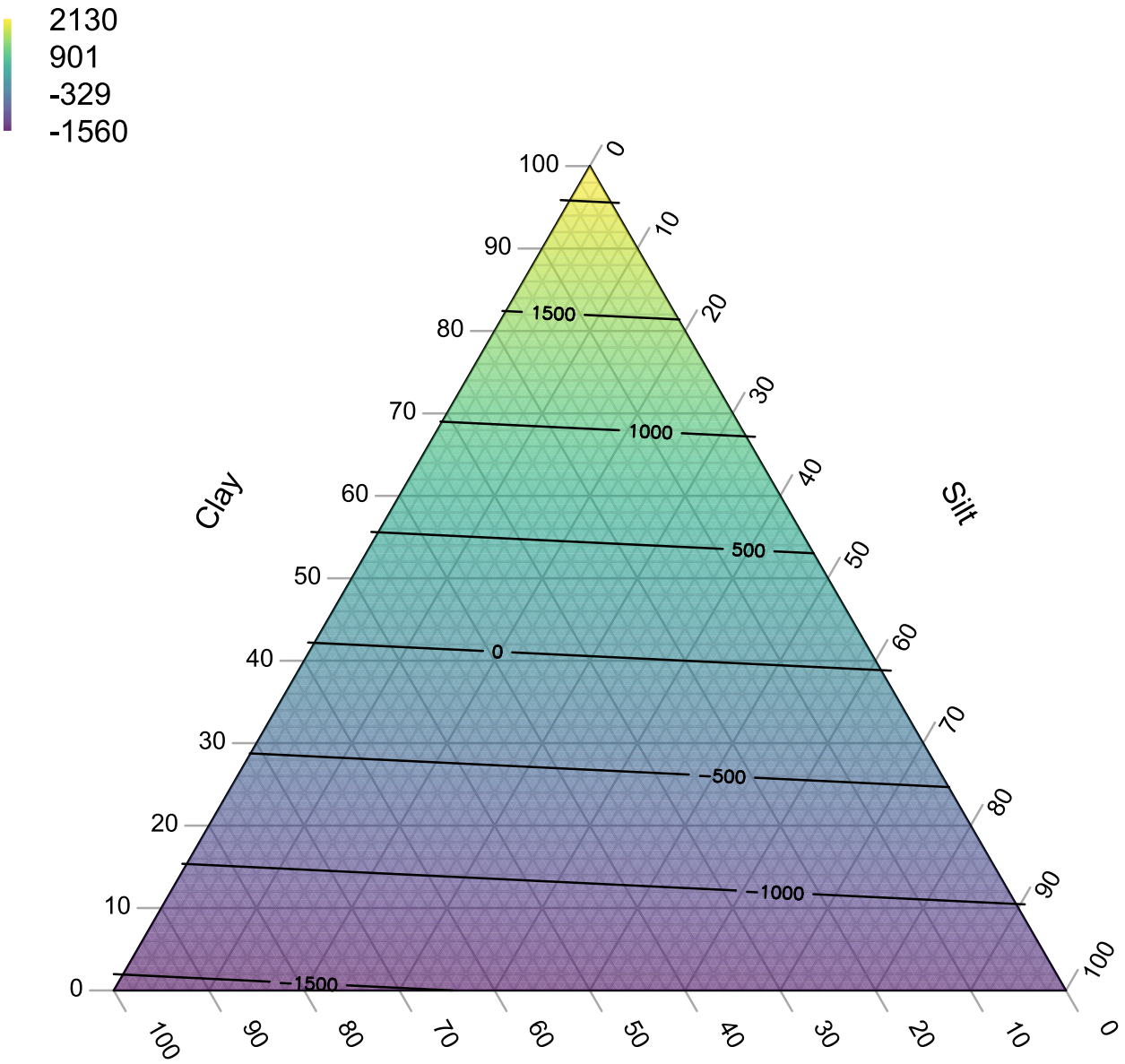


Figure 4.11. Soil texture triangle that describes the soybean yield starting conditions based on texture composition. The soybean yield distribution is based on the regression coefficients of the final soybean yield prediction model. The unit for the black lines is kg/ha.

4.7 Supplementary Information

Suppl. Table 2. FAME biomarkers organized by taxonomic microbial group designation and fatty acid structure. Biomarkers that were used in this study are highlighted. Furthermore, information regarding their number of observations and detection is provided.

Fatty Acid	Taxonomic Group		Biomarker	References	Detected	Nr. of obs.
Saturated	-		**14:0**	-	Yes	*n=153* (100 %)
	Gram (+); Gram (-); Fungi; Eukaryotes		**16:0 / 18:0**	Zelles, 1997; Kerger, 1986	Yes	*n=153* (100 %)
	Generally considered Bacteria		**15:0 / 17:0**	Federle, 1986; Tunlid et al., 1989; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
Mid-Chain Branched	Gram Positive (+) Bacteria	Actinomycetes (Actinobacteria)	**10Me16:0 / 10Me17:0 / 10Me18:0 / 10Me19:1w7c**	Fisher et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Forstegard et al., 1993, Veum et al. 2021	Yes	*n=153* (100 %)
Terminally Branched		Gram (+)	*i14:0 **	Zelles, 1999	Yes	*n=132* (86.3 %)
			a15:0 / i15:0 / i16:0 / a17:0 / i17:0	O'Leary and Wilkinson, 1988	Yes	*n=153 (100 %)*
Mid-Chain Branched / Monounsaturated	Gram Negative (-) Bacteria	*Sulfate red. Bacteria **	10Me16:0 / cy17:0 * <u>*anoxic and anaerobic conditions*</u>	Dowling et al., 1985, 1988; Parkes et al. 1993	-	-
Hydroxy-substituted		Gram (-)	2OH 12:0 / 3OH 12:0 / 2OH 14:0 / * 3OH 14:0 / 2OH 16:0 / 2OH 18:0	Parker et al., 1982	No	*n=0* (0 %)
			17:1 w8c	Zelles, 1999	Yes	*n=153 (100 %)*
Monounsaturated			**16:1ω7c / 18:1ω7c**	Wilkinson, 1988; Tunlid et al., 1989;	Yes	*n=153 (100 %)*
			cy17:0 / cy19:0	Wilkinson, 1988; Kerger, 1986	Yes	*n=153 (100 %)*
	Methanogens * (anaerobic conditions)		Type I: 16:1ω5t / 16:1ω7c / 16:1ω8c / 3OH 16:0 *	Nichols et al. 1985; Bowman et al. 1991, 1993	-	-
			Type II: 18:1ω8c / 18:1ω7c *		-	-

continues

Monounsaturated	Fungi	Fungi; *Plants* *	20:1ω9c *	Madan et al., 2002	Yes	*n=92* (60.1 %)
		Arbuscular Mycorrhizal Fungi (AMF); *Plants*	**16:1ω5c**	Olsson et al., 1995; Olsson, 1999; Madan et al., 2002	Yes	*n=153 (100 %)*
		Saprophytic Fungi; P*lants*	**18:1ω9c**	Vestal and White, 1989; Wallis et al. 2021	Yes	*n=153 (100 %)*
Polyunsaturated			**18:2ω6c**	Federle, 1986; Zelles, 1997; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
			18:3ω6c *	Federle, 1986; Klug, 1996	Yes	*n=140* (91.5 %)
		Fungi; *Plants* *	18:3ω3c *	Zelles, 1997	No	*n=0* (0 %)
		Fungi; *Plants* *	20:5ω3c *	Nordby et al., 1981; Olsson et al., 1995	Yes	*n=109* (71.2 %)
	Protozoa		**20:3ω6c**	Nordby et al., 1981; Guckert et al., 1985	Yes	*n=13* (8.5 %)
			20:4ω6c		Yes	*n=153* (100 %)
Saturated (Long Chain)	Eukaryotes		**21:0 / 22:0 / 24:0**	Zelles, 1999	Yes	*n=153* (100 %)
			23:0		Yes	*n=98* (64.1 %)

* Biomarkers were not classified in this study as such due to different environmental conditions or number of observations

Suppl. Table 3. Enzyme Activity Assay protocols for the individual enzymes (β-GLU, AS, NAG) when 1 g of soil is used.

Enzyme Assay Description and ecological role	Incubation step (37 °C for 1h)		After Incubation	
	Buffer	Substrate [a)]	$CaCl_2$ (0.5 M)	Stop solution
β-glucosidase (C cycling)	MUB pH 6.0; 4 mL	p-Nitrophenyl-β-D-glucopyranoside (0.05 M); 1 mL Sigma N7006	1 mL	THAM pH 12 (0.1 M); 4 mL
Arylsulfatase (S cycling)	0.5 M Acetate Buffer pH 5.8; 4 mL	p-Nitrophenyl sulfate (0.05 M); 1 mL; Sigma N3877	1 mL	THAM pH 12 (0.1 M); 4 mL
β-glucosaminidase (C and N cycling)	0.1 M Acetate Buffer pH 5.5; 4 mL	p-Nitrophenyl-N-acetyl-β-D-glucosaminide (0.01 M); 1 mL; Sigma N9376	1 mL	THAM pH 12 (0.1 M); 4 mL

a) Substrates were prepared by using the corresponding incubation buffer (Tabatabai, 1994).

Suppl. Table 4. List of all variables used in the multivariate optimized model including the corresponding regression coefficients. Each variable was placed into a predetermined group descriptor.

Basic Factors

		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Basic Factors	**Type**	Organic or Conventional Farming	-2899.762630	1 or 0
	Texture	Sand (0.05 - 0.002 mm)	-201.953858	%
		Silt (0.05 - 2 mm)	3525.740853	%
		Clay (<0.002 mm)	0	%

Tier 1 Soil Health Indicators

	Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Tier 1 Indicator (Core)	Soil Organic Carbon	782.859081	C %
	Total Nitrogen	3223.775072	N %

Soil Chemical Factors

		Classification		*Reg. Coef.* (β_P)	*Var. Unit*
Tier 1 Indicator (Chemical)	**pH**	pH		-561.551404	-
	pH Range	6.0 – 7.0		-44.7383024	1 or 0
		5.8 - 6.0	7.0 -7.4	0	1 or 0
		5.4 - 5.8	7.4 - 7.8	256.243034	1 or 0
		5.0 - 5.4	7.8 - 8.2	197.700105	1 or 0
		< 5.0	> 8.2	368.6723965	1 or 0

Land Management Factors

		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Agricultural Land Management Factors	**Cover Crop usage**	CC usage once in 3 years	447.324475	1 or 0
		CC usage twice in 3 years	294.184	1 or 0
		CC usage every years	136.62392	1 or 0
		No CC was planted	0	1 or 0
	Pesticide usage	No Herbicide (primarily org. Farming)	1741.8904	1 or 0
		Only Glyphosate	0	1 or 0
		Glyph. Mix	-394.55962	1 or 0
		Glyph. Mix + Dicamba	-563.507867	1 or 0
		Fungicide (Yes or No)	-157.22929	1 or 0
	Manure usage	Manure usage once in 3 years	666.340197	1 or 0
		Manure usage twice in 3 years	-11.737454	1 or 0
		Manure usage every years	-2793.7547	1 or 0
		No Manure was used	0	1 or 0
		Chicken Manure usage	136.036874	1 or 0
		Cattle Manure usage	2317.01089	1 or 0

continues

Environmental Factors

		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Environmental Factors	Growing Phase	Precipitation during growth season	908.9363344	mm
		Growing Time (planting to harvest)	-256.738567	days
	Soil Order	Mollisol or Alfisol	961.3298207	1 or 0

Enzymatic Factors

		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Enzyme	Enzyme Activity	GLU	0	mg PNP kg-1 soil h-1
		NAG	97.01485	
		AS	-24.81355	
		GLUAS	-82.64762	
		GLUNAG	0	
		NAGAS	0	
		GLUNAGAS	0	
	Enzyme Ratios	GLU / AS	25.25456	-
		GLU / NAG	-33.15335	-
		NAG / AS	0	-
		(GLU + AS) / NAG	95.10662	-
		(GLU + NAG) / AS	0	-
		(AS + NAG) / GLU	37.52667	-

Land Management Factors

		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
Agricultural Land Management Factors	Planting date	April date	46.3076002	1 or 0
		May date	-25.132364	1 or 0
		June date	0	1 or 0
	Crop Rotation	Corn-Soybean	-9.197E-10	1 or 0
		Corn-Soybean-Wheat	386.109727	1 or 0
	Tillage Practice	No-Tillage	0	1 or 0
		Chisel Tillage	360.329681	1 or 0
		Chisel + Field Cultivator	65.0503949	1 or 0
		Disk Tillage	-176.45816	1 or 0
		Moldboard Plow	-260.56456	1 or 0
	Surface Residue Coverage	Residue Coverage	1026.79374	%
		> 60% Coverage	-187.39399	1 or 0
		30 - 60 % Coverage	10.7071975	1 or 0
		15 - 30 % Coverage	0	1 or 0
		6 - 15 % Coverage	194.79937	1 or 0
		< 6% Coverage	-638.59513	1 or 0

continues

Soil Biological Factors				
		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
EL-FAME	FAME Biomarkers	15:0	-3.564598731	nmol / g- dry soil
		17:0	-479.0980938	
		16:0 10-methyl	319.9718984	
		17:0 10-methyl	-710.5528935	
		18:0 10-methyl	1293.0878	
		15:0 iso	-2975.468094	
		15:0 anteiso	2311.285751	
		16:0 iso	-1639.354119	
		17:0 iso	0	
		17:0 anteiso	0	
		16:1 w7c	1257.409324	
		17:0 cyclo w7c	3484.17086	
		19:0 cyclo w7c	-3.114772268	
		18:1 w7c	-50.57138302	
		16:1 w5c	0	
		18:2 w6c	-173.5713654	
		18:1 w9c	445.2915434	
		20:4 w6c	-800.8416487	
		14:0	-1667.600833	
		16:0	298.3428361	
		18:0	602.7170674	
		20:0	-898.5643857	
		17:1 w8c	-1912.408918	
		17:1 w7c 10-methyl	-18.08884181	
		19:1 w7c 10-methyl	0	
		new 21:0	540.1253847	
		21:0	-6157.119087	
		22:0	-7158.579225	
		22:0 iso	316.7550042	
		24:0	-4265.224602	
		Total Biomarker	0	

Soil Biological Factors				
		Classification	*Reg. Coef.* (β_P)	*Var. Unit*
EL-FAME	FAME Biomarkers	Actinobacteria	0	nmol / g- dry soil
		Eukaryotes	15056.97471	
		Fungi	-4.517994941	
		Arbuscular Mycorrhizal Fungi	0	
		Gram- bacteria	1123.848558	
		Gram+ bacteria	-854.2835212	
		Protozoa	0	
		Total Biomarker	0	
		Total Fungi	0	

TRIPLETT-VAN DOREN LONG-TERM NO-TILL PLOTS **(2021)**

SNYDER FARM (WOOSTER, OH): 731, 732, AND 741

GRAVEL ROAD LEADING TO FARM BUILDINGS

731

PLOT	TILL	ROT.	CROP
101	CH	CAA	ALFALFA 2
102	CH	CAA	ALFALFA 1
103	CH	CAA	CORN
104	CH	CS	CORN
105	CH	CS	SOYBEAN
106	CH	CC	CORN
107	NT	CAA	ALFALFA 1
108	NT	CAA	CORN
109	NT	CAA	ALFALFA 2
201	NT	CS	CORN
202	NT	CS	SOYBEAN
203	NT	CAA	ALFALFA 2
204	NT	CAA	CORN
205	NT	CAA	ALFALFA 1
206	NT	CC	CORN
207	PL	CAA	CORN
208	PL	CAA	ALFALFA 2
209	PL	CAA	ALFALFA 1
301	PL	CS	SOYBEAN
302	PL	CS	CORN
303	PL	CC	CORN
304	PL	CAA	ALFALFA 2
305	PL	CAA	ALFALFA 1
306	PL	CAA	CORN
307	CH	CC	CORN
308	CH	CS	CORN
309	CH	CS	SOYBEAN

732

PLOT	TILL	ROT.	CROP
110	PL	CC	CORN
111	PL	CAA	ALFALFA 1
112	PL	CAA	CORN
113	PL	CAA	ALFALFA 2
114	PL	CS	SOYBEAN
115	PL	CS	CORN
116	NT	CS	SOYBEAN
117	NT	CS	CORN
118	NT	CC	CORN
210	CH	CS	SOYBEAN
211	CH	CS	CORN
212	CH	CC	CORN
213	CH	CAA	CORN
214	CH	CAA	ALFALFA 2
215	CH	CAA	ALFALFA 1
216	PL	CC	CORN
217	PL	CS	CORN
218	PL	CS	SOYBEAN
310	NT	CC	CORN
311	NT	CAA	CORN
312	NT	CAA	ALFALFA 1
313	NT	CAA	ALFALFA 2
314	NT	CS	SOYBEAN
315	NT	CS	CORN
316	CH	CAA	ALFALFA 1
317	CH	CAA	CORN
318	CH	CAA	ALFALFA 2

N

741

CC

PLOT	TILL
401	NT
402	CH
403	PL
501	CH
502	NT
503	PL
601	NT
602	CH
603	PL
701	PL
702	NT
703	CH

Plot Dimensions

731/732:

75' L x 14' W

741:

120' L x 21' W

KEY:

CC = CONTINUOUS CORN

CS = CORN/SOYBEAN

CAA = CORN/ALFALFA 1/ALFALFA 2

NT = NO TILL

CH = CHISEL

PL = PLOW (MOLDBOARD)

Suppl. Figure 1. 2021 Plot map for the Triplett-Van Doren Long-Term Field Site in Wooster, Ohio.

LONG-TERM TILLAGE PLOTS (2021)
NORTHWEST RESEARCH STATION (Hoytville, OH): FieldTA - 3

N

PLOT	TILL	ROT.	CROP
101	CH	CAA	CORN
102	CH	CAA	ALFALFA 2
103	CH	CAA	ALFALFA 1
104	NT	CS	SOYBEAN
105	NT	CS	CORN
106	PL	CAA	CORN
107	PL	CAA	ALFALFA 2
108	PL	CAA	ALFALFA 1
109	PL	CC	CORN
110	CH	CS	SOYBEAN
111	CH	CS	CORN
112	NT	CC	CORN
113	NT	CAA	CORN
114	NT	CAA	ALFALFA 2
115	NT	CAA	ALFALFA 1
116	PL	CS	SOYBEAN
117	PL	CS	CORN
118	CH	CC	CORN

PLOT	TILL	ROT.	CROP
201	PL	CS	CORN
202	PL	CS	SOYBEAN
203	CH	CC	CORN
204	NT	CAA	ALFALFA 2
205	NT	CAA	CORN
206	NT	CAA	ALFALFA 1
207	PL	CC	CORN
208	CH	CS	CORN
209	CH	CS	SOYBEAN
210	CH	CAA	ALFALFA 2
211	CH	CAA	CORN
212	CH	CAA	ALFALFA 1
213	NT	CC	CORN
214	NT	CS	CORN
215	NT	CS	SOYBEAN
216	PL	CAA	ALFALFA 2
217	PL	CAA	CORN
218	PL	CAA	ALFALFA 1

PLOT	TILL	ROT.	CROP
301	NT	CC	CORN
302	PL	CC	CORN
303	NT	CS	SOYBEAN
304	NT	CS	CORN
305	CH	CAA	ALFALFA 1
306	CH	CAA	CORN
307	CH	CAA	ALFALFA 2
308	PL	CS	SOYBEAN
309	PL	CS	CORN
310	CH	CC	CORN
311	PL	CAA	ALFALFA 1
312	PL	CAA	CORN
313	PL	CAA	ALFALFA 2
314	CH	CS	SOYBEAN
315	CH	CS	CORN
316	NT	CAA	ALFALFA 1
317	NT	CAA	CORN
318	NT	CAA	ALFALFA 2

KEY:

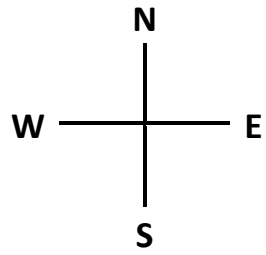
CC = CONTINUOUS CORN
CS = CORN/SOYBEAN
CAA = CORN/ALFALFA 1/ALFALFA 2

NT = NO TILL
CH = CHISEL
PL = PLOW (MOLDBOARD)

PLOTS: 20' W x 90' L

Suppl. Figure 2. 2021 Plot map for the Northwest Research Station in Hoytville, Ohio.

East Straw Mulch Experiment



Rep 1	**16** **M8F0**	**13** **M0F1**	**10** **M16F1**	**7** **M8F1**	**4** **M0F0**	**1** **M16F0**
Rep 2	**17** **M0F1**	**14** **M8F0**	**11** **M0F0**	**8** **M16F1**	**5** **M16F0**	**2** **M8F1**
Rep 3	**18** **M16F0**	**15** **M0F0**	**12** **M8F0**	**9** **M8F1**	**6** **M16F1**	**3** **M0F1**

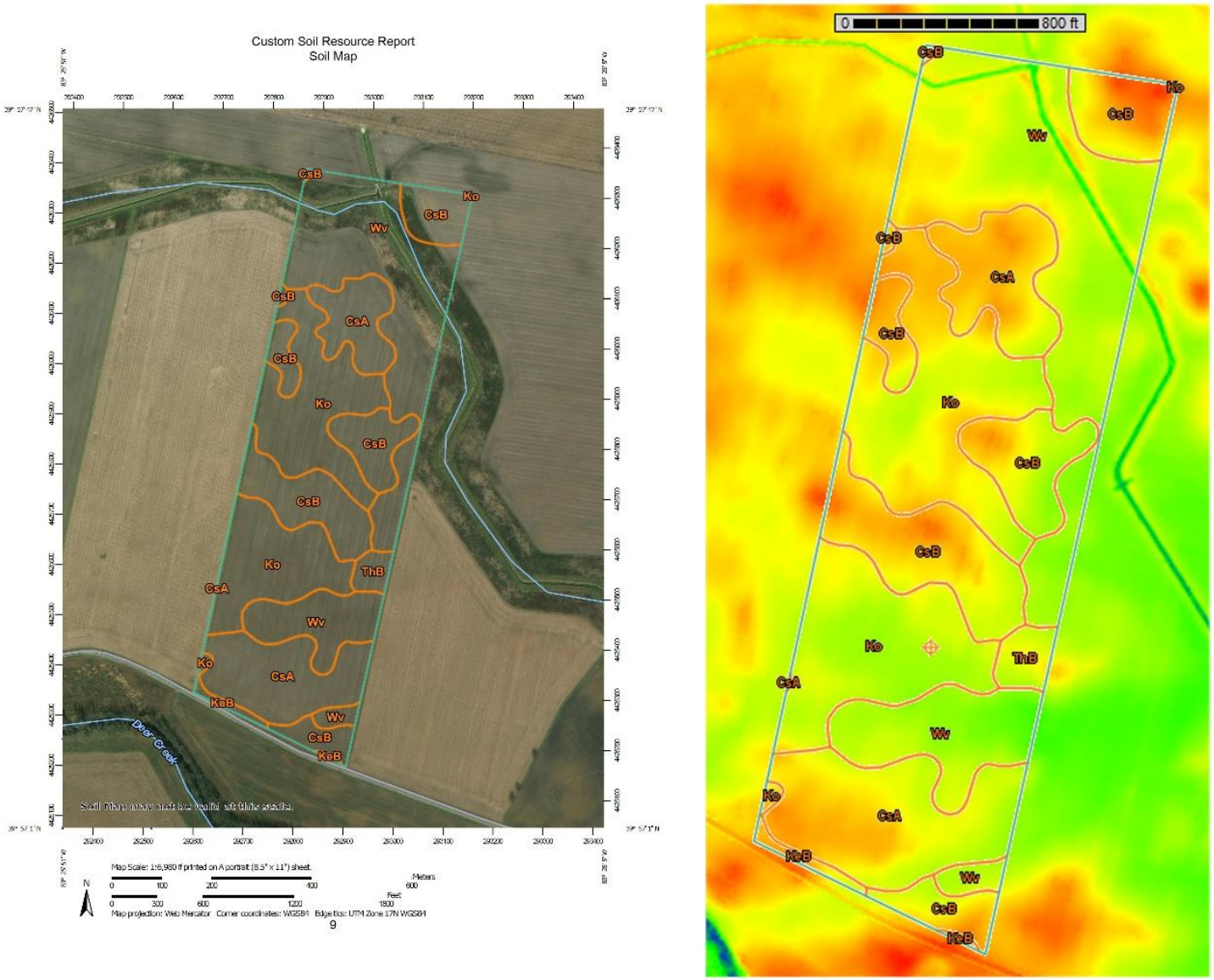
5m (16.4 ft)

5m (16.4 ft)

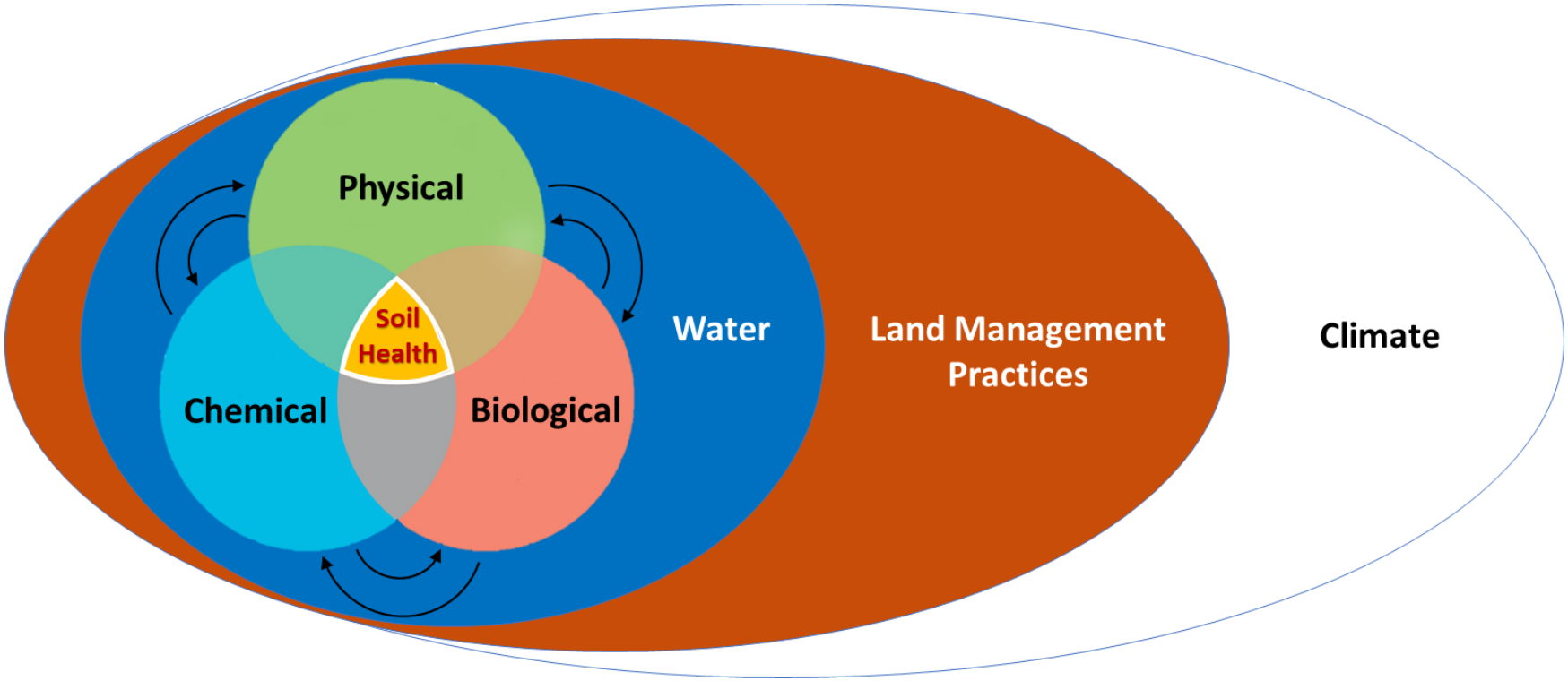
Fertilizer
F0 - No Fertilizer
F1 - 244 kg/ha N (184kg as Urea, 60 kg as NPK)

Straw Mulch
M0 - No Mulch
M8 - 8 Mg/ha mulch (20kg/plot)
M16 - 16 Mg/ha mulch (40kg/plot)

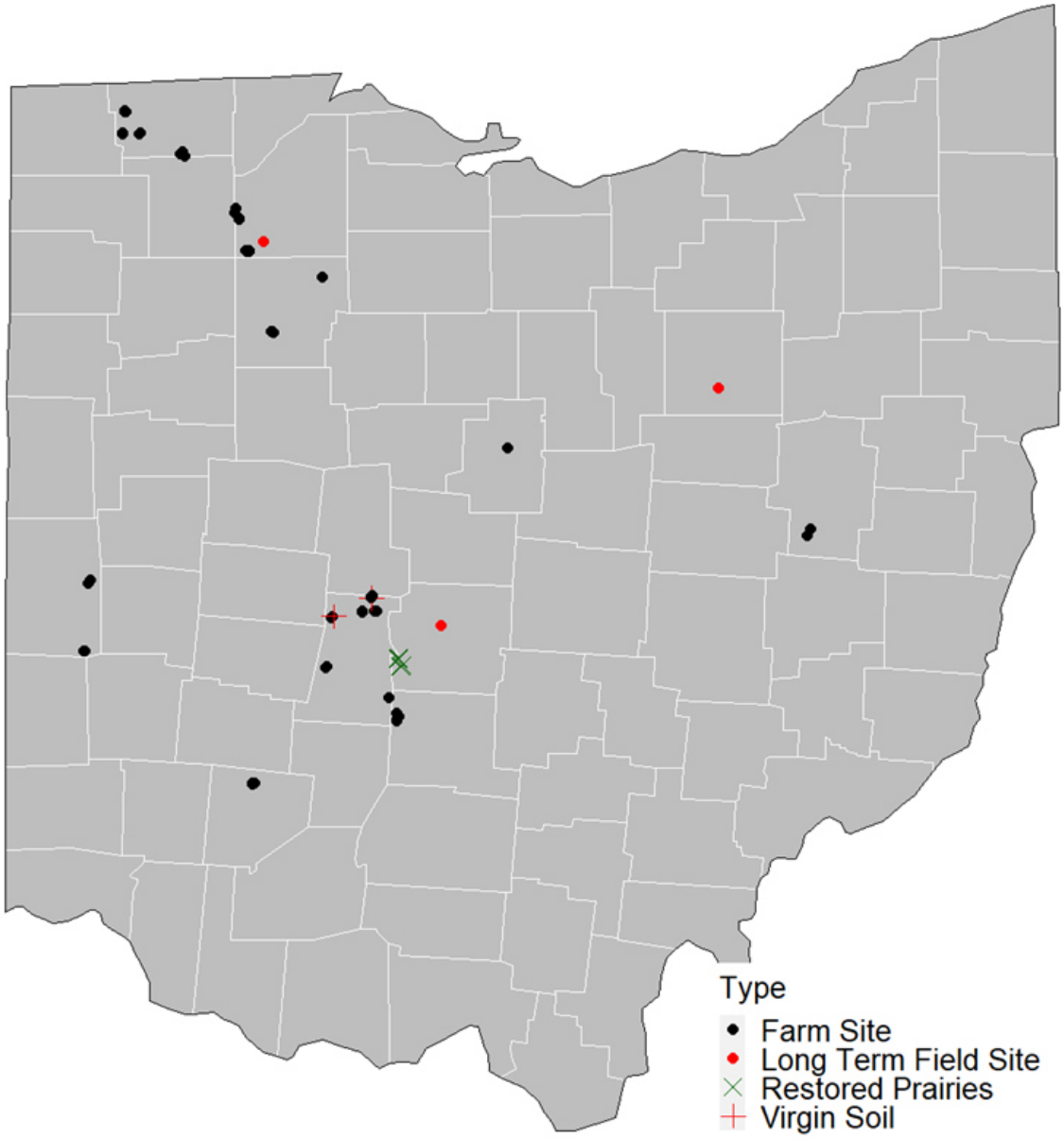
Suppl. Figure 3. 2021 Plot map for the East Straw Mulch Experiment in Columbus, Ohio.



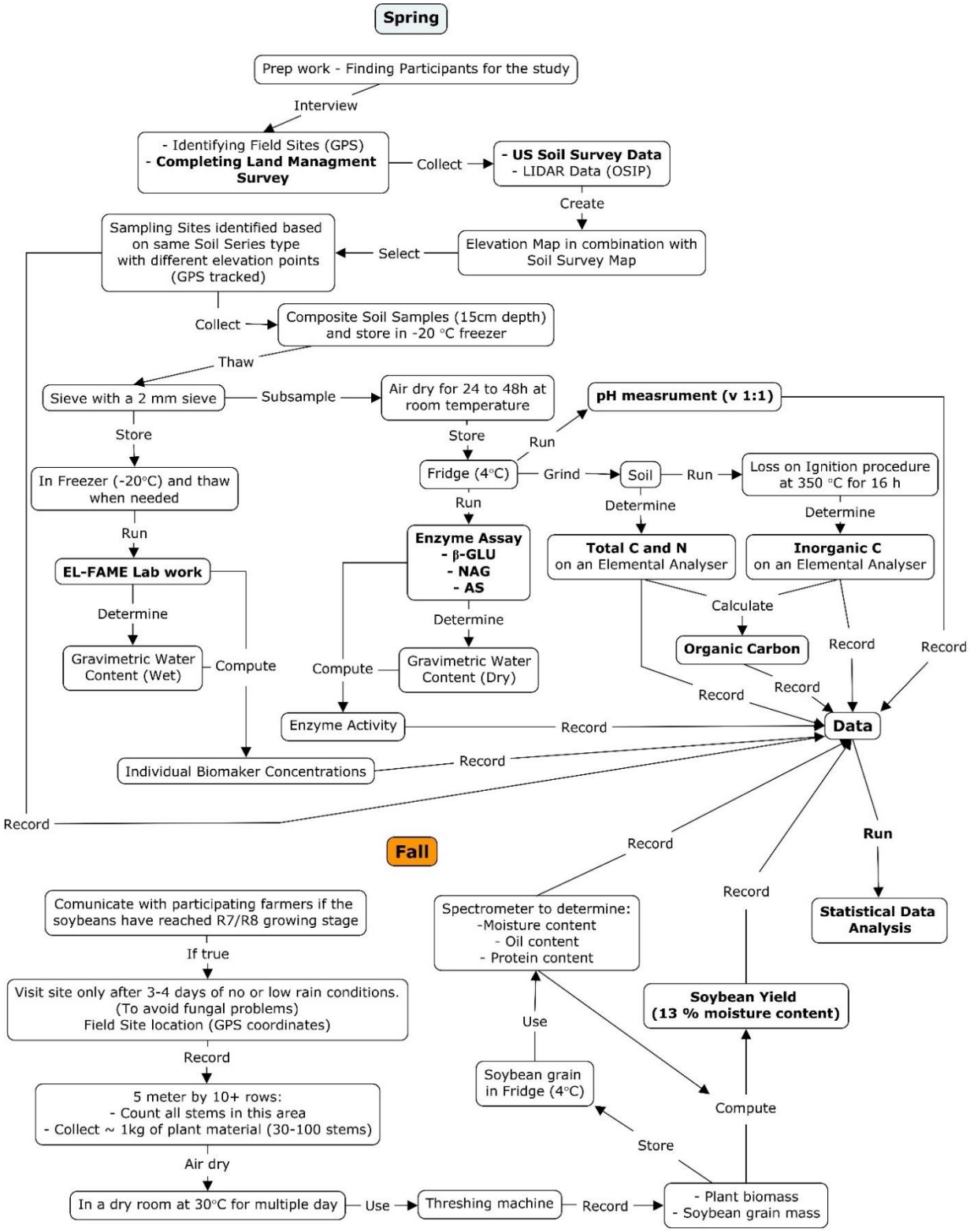
Suppl. Figure 4. Example of a soil map overlayed on top of an elevation heat map created with LIDAR data



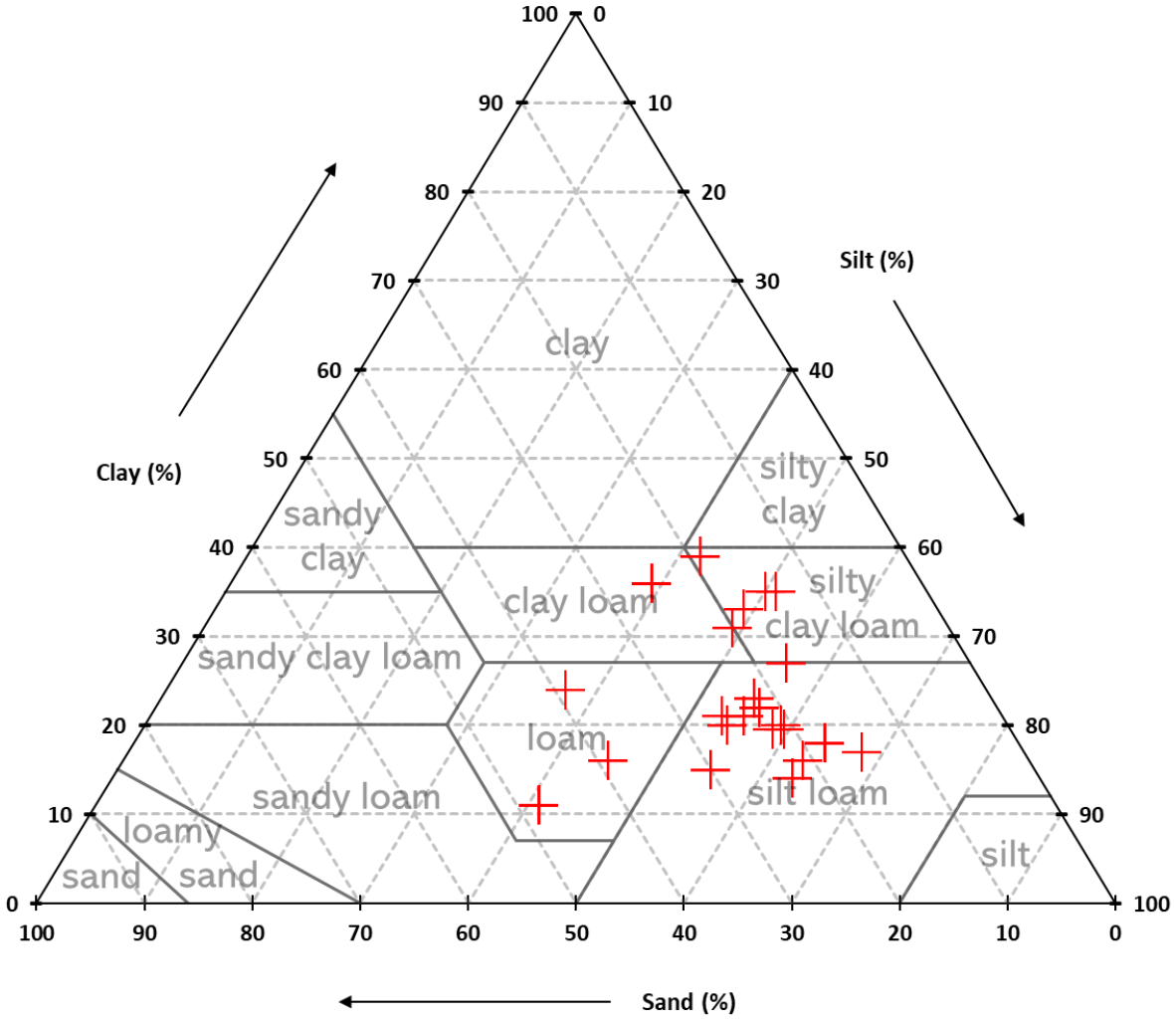
Suppl. Figure 5. Simplified representation of the interactions between physical, chemical, and biological factors which influence and define Soil Health including the overreaching factors that can influence soil health indicators.



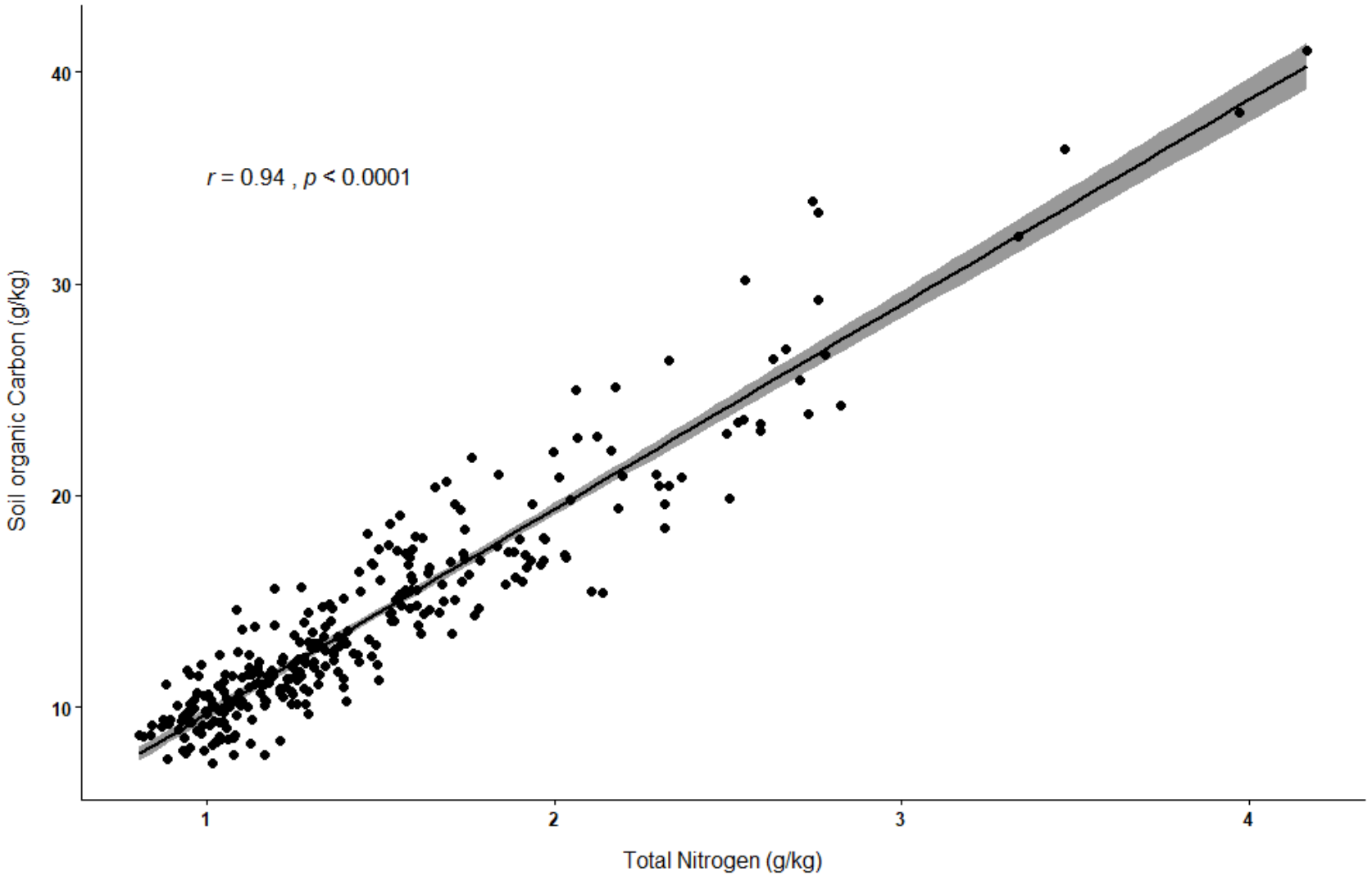
Suppl. Figure 6. Ohio map depicting all study site locations separated into farm (black circle), long-term field (red circle), restored prairie (green X symbol), and virgin soil sites (red cross). The Kellogg Long-Term Ecological Research site located in Michigan is not depicted.



Suppl. Figure 7. Flow Chart of the experimental study design representing soil related measurements in spring and soybean yield related measurement steps in fall.



Suppl. Figure 8. Soil texture triangle with all soil types (red cross) used in this study.



Suppl. Figure 9. Scatter plot for Soil org. Carbon and Total Nitrogen (n=301).

5. Appendix

Appendix Tab.1) List of all agricultural farm sites in the study. Separated into organic and conventional farm types. Basic location description, soil organic carbon, planted crop, soil texture, and soil series information is included.

ID	Farming Type	Site ID	Field ID	Elevation	State	County	Growing Period (days)			Planted Crop			Soil Texture			Soil Series
							year 1	year 2	year 3	year 1	year 2	year 3	Clay (%)	Silt (%)	Sand (%)	
1	Org	VY	F1	HE	OH	Madison	123	-	-	Soy	Wheat	Corn	18	64	18	Crosby-Lewisburg silt loams
2	Org	VY		LE	OH	Madison	123	-	-	Soy	Wheat	Corn	18	64	18	
3	Org	VY	F2	HE	OH	Madison	-	137	-	Corn	Soy	Wheat	18	64	18	Crosby-Lewisburg silt loams
4	Org	VY		LE	OH	Madison	-	137	-	Corn	Soy	Wheat	18	64	18	
5	Org	LM	F1	HE	OH	Madison	122	-	-	Soy	Wheat	Corn	18	64	18	Crosby-Lewisburg silt loams
6	Org	LM		LE	OH	Madison	122	-	-	Soy	Wheat	Corn	18	64	18	
7	Org	LM	F2	HE	OH	Madison	-	128	-	Corn	Soy	Wheat	18	64	18	Crosby-Lewisburg silt loams
8	Org	LM		LE	OH	Madison	-	128	-	Corn	Soy	Wheat	18	64	18	
9	Org	JK	F1	HE	OH	Madison	120	-	-	Soy	Wheat	Corn	22	58.5	19.5	Crosby-Lewisburg silt loams
10	Org	JK		LE	OH	Madison	120	-	-	Soy	Wheat	Corn	18	64	18	
11	Org	JK	F2	HE	OH	Madison	-	-	132	Wheat	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
12	Org	JK		LE	OH	Madison	-	125	-	Corn	Soy	Wheat	18	64	18	
13	Org	DB	F1	HE	OH	Hancock	160	-	-	Soy	Corn	Wheat	15	50	35	Pewamo silty clay loam
14	Org	DB		LE	OH	Hancock	160	-	-	Soy	Corn	Wheat	22	56	22	Blount silt loam
15	Org	DB	F2	HE	OH	Hancock	-	132	-	Corn	Soy	Wheat	39	37	24	Blount-Houcktown complex
16	Org	DB		LE	OH	Hancock	-	132	-	Corn	Soy	Wheat	22	56	22	Blount silt loam

continues

17	Org	RA	F1	HE	OH	Clinton	-	131	-	Oats / Forage	Soy	Hay / Alfalfa	17	56	27	Treaty silty clay loam
18	Org	RA		LE	OH	Clinton	-	131	-	Oats / Forage	Soy	Hay / Alfalfa	17	56	27	
19	Org	RA	F2	HE	OH	Clinton	-	-	-	Soy	Oats / Forage	Forage	17	56	27	Treaty silty clay loam
20	Org	RA		LE	OH	Clinton	-	-	-	Soy	Oats / Forage	Forage	17	56	27	
21	Conv	BH	F1	HE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	Hoytville clay loam
22	Conv	BH		LE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	
23	Conv	BH	F2	HE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	Hoytville clay loam
24	Conv	BH		LE	OH	Fulton	-	-	141	Bare	Corn	Soy	25	39	36	
25	Conv	PD	F1	HE	OH	Fulton	-	156	-	Rye	Soy	Corn	14	51	35	Lenawee silty clay loam
26	Conv	PD		LE	OH	Fulton	141	-	151	Soy	Corn	Soy	14	51	35	
27	Conv	PD	F2	HE	OH	Fulton	134	149	-	Soy	Soy	Corn	18	49	33	Fulton silty clay loam
28	Conv	PD		LE	OH	Fulton	134	149	-	Soy	Soy	Corn	18	49	33	
29	Conv	PD	F3	HE	OH	Fulton	118	-	-	Soy	Bare	Corn	25	39	36	Hoytville clay loam
30	Conv	PD		LE	OH	Fulton	118	133	-	Soy	Soy	Corn	25	39	36	
31	Conv	PD	F4	HE	OH	Fulton	132	136	-	Soy	Soy	Corn	25	39	36	Hoytville clay loam
32	Conv	PD		LE	OH	Fulton	132	136	-	Soy	Soy	Corn	25	39	36	
33	Conv	PD	F5	HE	OH	Fulton	135	-	129	Soy	Corn	Soy	22	55	23	Del Rey silt loam
34	Conv	PD		LE	OH	Fulton	135	-	129	Soy	Corn	Soy	22	55	23	
35	Conv	PD	F6	HE	OH	Fulton	121	132	-	Soy	Soy	Corn	25	39	36	Hoytville clay loam
36	Conv	PD		LE	OH	Fulton	121	-	136	Soy	Corn	Soy	25	39	36	

continues

37	Conv	PD	F7	HE	OH	Fulton	142	-	130	Soy	Corn	Soy	22	55	23	Del Rey silt loam
38	Conv	PD		LE	OH	Fulton	142	-	130	Soy	Corn	Soy	22	55	23	
39	Conv	TL	F1	HE	OH	Morrow	134	138	-	Soy	Soy	Corn	21	63	16	Centerburg silt loam
40	Conv	TL		LE	OH	Morrow	134	138	-	Soy	Soy	Corn	21	59	20	
41	Conv	TL	F2	HE	OH	Morrow	134	-	139	Soy	Corn	Soy	21	63	16	Centerburg silt loam
42	Conv	TL		LE	OH	Morrow	134	-	139	Soy	Corn	Soy	21	63	16	
43	Conv	JM	F1	HE	OH	Hancock	-	-	152	Bare	Corn	Soy	22	56	22	Blount silt loam, ground moraine
44	Conv	JM		LE	OH	Hancock	-	-	152	Bare	Corn	Soy	22	56	22	
45	Conv	JM	F2	HE	OH	Hancock	-	140	-	Bare	Soy	Corn	22	56	22	Blount silt loam, ground moraine
46	Conv	JM		LE	OH	Hancock	-	140	-	Bare	Soy	Corn	22	56	22	
47	Conv	BG	F1	HE	OH	Hancock	-	168	-	Bare	Soy	Corn	19	42	39	Hoytville clay loam
48	Conv	BG		LE	OH	Hancock	-	168	-	Bare	Soy	Corn	19	42	39	
49	Conv	BG	F2	HE	OH	Hancock	-	-	154	Bare	Corn	Soy	19	42	39	Hoytville clay loam
50	Conv	BG		LE	OH	Hancock	-	-	154	Bare	Corn	Soy	19	42	39	
51	Conv	RB	F1	HE	OH	Darke	126	163	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
52	Conv	RB		LE	OH	Darke	126	163	-	Soy	Soy	Corn	18	64	18	
53	Conv	AO	F1	HE	OH	Darke	120	123	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
54	Conv	AO		LE	OH	Darke	120	123	-	Soy	Soy	Corn	18	64	18	
55	Conv	AO	F2	HE	OH	Darke	114	-	145	Soy	Corn	Soy	18	64	18	Crosby silt loam
56	Conv	AO		LE	OH	Darke	114	-	145	Soy	Corn	Soy	18	64	18	

continues

57	Conv	BM	F1	HE	OH	Madison	156	-	151	Soy	Corn	Soy	24	55	21	Miamian-Lewisburg silt loams
58	Conv	BM		LE	OH	Madison	156	-	151	Soy	Corn	Soy	24	55	21	
59	Conv	BM	F2	HE	OH	Madison	-	164	-	Corn	Soy	Corn	24	55	21	Miamian-Lewisburg silt loams
60	Conv	BM		LE	OH	Madison	-	164	-	Corn	Soy	Corn	24	55	21	
61	Conv	CO	F1	HE	OH	Pickaway	123	135	-	Soy	Soy	Corn	18	64	18	Crosby silt loam
62	Conv	CO		LE	OH	Pickaway	123	135	-	Soy	Soy	Corn	18	64	18	
63	Conv	CO	F2	HE	OH	Pickaway	-	126	145	Corn	Soy	Soy	18	64	18	Crosby silt loam
64	Conv	CO		LE	OH	Pickaway	-	126	145	Corn	Soy	Soy	18	64	18	
65	Conv	CHA	F1	HE	OH	Pickaway	-	134	-	Corn	Soy	Corn	26	53	21	Miamian-Lewisburg silt loams
66	Conv	CHA		LE	OH	Pickaway	-	134	-	Corn	Soy	Corn	26	53	21	
67	Conv	CHA	F2	HE	OH	Pickaway	-	134	-	Corn	Soy	Corn	21	59.5	19.5	Crosby silt loam
68	Conv	CHA		LE	OH	Pickaway	-	134	-	Corn	Soy	Corn	18	64	18	
69	Conv	LM	F1	HE	OH	Tuscarawas	173	-	140	Soy	Corn	Soy	39	45	16	Wheeling loam
70	Conv	LM		LE	OH	Tuscarawas	173	-	140	Soy	Corn	Soy	39	45	16	
71	Conv	LM	F2	HE	OH	Tuscarawas	-	131	-	Corn	Soy	Corn	39	45	16	Wheeling loam
72	Conv	LM		LE	OH	Tuscarawas	-	131	-	Corn	Soy	Corn	39	45	16	
73	Conv	JH	F1	HE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	Crosby-Lewisburg silt loams
74	Conv	JH		LE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	
75	Conv	JH	F1 - Tra	HE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	Crosby-Lewisburg silt loams
76	Conv	JH		LE	OH	Madison	-	138	-	Corn	Soy	Corn	18	64	18	

continues

77	Conv	JH		HE1	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	
78	Conv	JH	F2	HE2	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
79	Conv	JH		LE	OH	Madison	-	-	138	Soy	Corn	Soy	18	64	18	
80	Conv	BP	F1	HE	OH	Madison	-	-	154	Soy	Corn	Soy	18	64	18	Crosby-Lewisburg silt loams
81	Conv	BP		LE	OH	Madison	-	-	154	Soy	Corn	Soy	18	64	18	
						Mean:	133	140	143				21.3	55.2	23.5	

Appendix Tab.2) List of FAME biomarkers organized by taxonomic microbial group designation and fatty acid structure. Biomarkers that were used in this study are highlighted. Furthermore, information regarding their number of observations and detection is provided.

Fatty Acid	Taxonomic Group		Biomarker	References	Detected	Nr. of obs.
Saturated	-		**14:0**	-	Yes	*n=153* (100 %)
	Gram (+); Gram (-); Fungi; Eukaryotes		**16:0 / 18:0**	Zelles, 1997; Kerger, 1986	Yes	*n=153* (100 %)
	Generally considered Bacteria		**15:0 / 17:0**	Federle, 1986; Tunlid et al., 1989; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
Mid-Chain Branched	Gram Positive (+) Bacteria	Actinomycetes (Actinobacteria)	**10Me16:0 / 10Me17:0 / 10Me18:0 / 10Me19:1w7c**	Fisher et al., 1983; Kroppenstedt, 1985; Zelles, 1997; Forstegard et al., 1993, Veum et al. 2021	Yes	*n=153* (100 %)
Terminally Branched		Gram (+)	*i14:0 **	Zelles, 1999	Yes	*n=132* (86.3 %)
			a15:0 / i15:0 / i16:0 / a17:0 / i17:0	O'Leary and Wilkinson, 1988	Yes	*n=153 (100 %)*
Mid-Chain Branched / Monounsaturated	Gram Negative (-) Bacteria	*Sulfate red. Bacteria **	10Me16:0 / cy17:0 * <u>*anoxic and anaerobic conditions*</u>	Dowling et al., 1985, 1988; Parkes et al. 1993	-	-
Hydroxy-substituted		Gram (-)	2OH 12:0 / 3OH 12:0 / 2OH 14:0 / * 3OH 14:0 / 2OH 16:0 / 2OH 18:0	Parker et al., 1982	No	*n=0* (0 %)
Monounsaturated			**17:1 w8c**	Zelles, 1999	Yes	*n=153 (100 %)*
			16:1ω7c / 18:1ω7c	Wilkinson, 1988; Tunlid et al., 1989;	Yes	*n=153 (100 %)*
			cy17:0 / cy19:0	Wilkinson, 1988; Kerger, 1986	Yes	*n=153 (100 %)*
	Methanogens * (anaerobic conditions)		Type I: 16:1ω5t / 16:1ω7c / 16:1ω8c / 3OH 16:0 *	Nichols et al. 1985; Bowman et al. 1991, 1993	-	-
			Type II: 18:1ω8c / 18:1ω7c *		-	-

continues

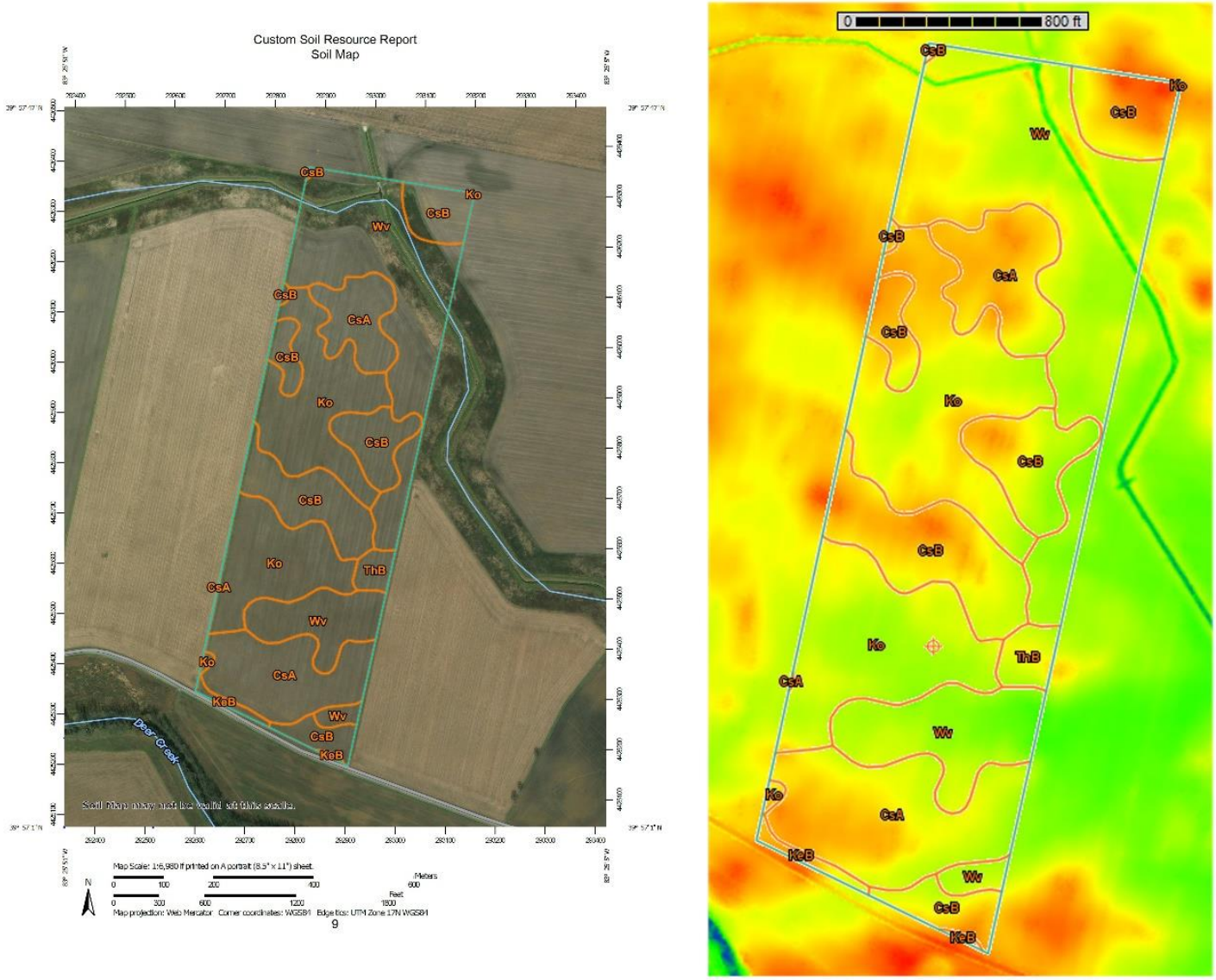
Monounsaturated	Fungi	Fungi; *Plants* *	20:1ω9c *	Madan et al., 2002	Yes	*n=92* (60.1 %)
		Arbuscular Mycorrhizal Fungi (AMF); *Plants*	**16:1ω5c**	Olsson et al., 1995; Olsson, 1999; Madan et al., 2002	Yes	*n=153 (100 %)*
Polyunsaturated		Saprophytic Fungi; P*lants*	**18:1ω9c**	Vestal and White, 1989; Wallis et al. 2021	Yes	*n=153 (100 %)*
			18:2ω6c	Federle, 1986; Zelles, 1997; Forstegard and Baath, 1996	Yes	*n=153* (100 %)
			18:3ω6c *	Federle, 1986; Klug, 1996	Yes	*n=140* (91.5 %)
		Fungi; *Plants* *	18:3ω3c *	Zelles, 1997	No	*n=0* (0 %)
		Fungi; *Plants* *	20:5ω3c *	Nordby et al., 1981; Olsson et al., 1995	Yes	*n=109* (71.2 %)
	Protozoa		**20:3ω6c**	Nordby et al., 1981; Guckert et al., 1985	Yes	*n=13* (8.5 %)
			20:4ω6c		Yes	*n=153* (100 %)
Saturated (Long Chain)	Eukaryotes		**21:0 / 22:0 / 24:0**	Zelles, 1999	Yes	*n=153* (100 %)
			23:0		Yes	*n=98* (64.1 %)

* Biomarkers were not classified in this study as such due to different environmental conditions or number of observations

Appendix Tab.3) Enzyme Activity Assay protocols for the individual enzymes (β-GLU, AS, NAG) when 1g of soil is used.

Enzyme Assay Description and ecological role	Incubation step (37 °C for 1h)		After Incubation	
	Buffer	Substrate [a)]	$CaCl_2$ (0.5 M)	Stop solution
β-glucosidase (C cycling)	MUB pH 6.0; 4 mL	p-Nitrophenyl-β-D-glucopyranoside (0.05 M); 1 mL Sigma N7006	1 mL	THAM pH 12 (0.1 M); 4 mL
Arylsulfatase (S cycling)	0.5 M Acetate Buffer pH 5.8; 4 mL	p-Nitrophenyl sulfate (0.05 M); 1 mL; Sigma N3877	1 mL	THAM pH 12 (0.1 M); 4 mL
β-glucosaminidase (C and N cycling)	0.1 M Acetate Buffer pH 5.5; 4 mL	p-Nitrophenyl-N-acetyl-β-D-glucosaminide (0.01 M); 1 mL; Sigma N9376	1 mL	THAM pH 12 (0.1 M); 4 mL

a) Substrates were prepared by using the corresponding incubation buffer (Tabatabai, 1994).



Appendix Fig.1 Example of a soil map overlayed on top of an elevation heat map created with LIDAR data.