Potential of Soil and Crop Spatial Properties for Depicting Within-Field Nutrient Variability and Delineating Management Zones for Precision Soil Sampling

by

J.D. Beasley

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Crop, Soil, and Environmental Science, Master of Science

Auburn, Alabama
May 2nd, 2026

Keywords: nutrient management, management zones,
soil sampling, precision agriculture

Approved by

Dr. Simerjeet Virk, Chair, Associate Professor, Department of Biosystems Engineering
Dr. Audrey Gamble, Associate Professor, Department of Crop, Soils & Environmental Sciences
Dr. Thorsten J. Knappenberger, Associate Professor, Department of Crop, Soils & Environmental Sciences

Abstract

Limited information currently exists on the potential of various soil and crop spatial properties for delineating management zones (MZs) for precision soil sampling in the southeastern United States; therefore, two studies were conducted to evaluate the effectiveness of different spatial data layers for characterizing within-field nutrient variability and generating MZs for zone-based soil sampling as a cost-effective alternative to high-density grid sampling. The first study assessed relationships between spatial data layers including shallow and deep soil electrical conductivity (EC), elevation, normalized difference vegetation index (NDVI), normalized yield, and soil survey data, and soil pH, phosphorus (P), and potassium (K) variability. This study was performed across nine fields in Alabama and Georgia ranging in size from 9.6 to 37.8 ha. High-density grid soil samples (0.1 ha) were collected across all fields to establish baseline nutrient variability. Pearson correlation analysis and Random Forest (RF) modeling were performed to assess relationships between spatial layers and soil nutrient levels on a field-by-field basis. Soil P and K exhibited considerable spatial variability across all nine fields, with CV values reaching up to 46% and 39%, respectively, while soil pH showed comparatively little variability, with CV values not exceeding 7%. No single spatial layer showed consistent relationships across all fields; however, deep EC demonstrated significant correlations with all three nutrients in seven or more fields, and elevation showed significant associations across five to seven fields depending on the nutrient, while soil series, NDVI, and yield showed more limited and field-specific performance, with soil survey data demonstrating the weakest

relationships between all three nutrients. RF modeling resulted in R^2 values between 0.04 and 0.78, with soil K models showing the greatest consistency and soil pH models performing the weakest, reflecting the low inherent spatial variability of pH across the study fields. The second study evaluated four spatial data layer combinations for MZ delineation across four fields (11.8–30.1 ha) in the Coastal Plain region of Georgia, using k-means clustering at k = 3, 4, and 5. Delineation strategies included EC Only (S1), EC + Elevation (S2), EC + Elevation + NDVI (S3), and EC + Elevation + Yield (S4), with zone performance assessed via variance reduction (VR) of soil pH, P, and K relative to the 0.1-ha grid soil sampling baseline. Results indicated that S1 performed worst across all fields and zone counts, while S2 showed competitive performance with S3 and S4 across most fields. S3 and S4 produced notably higher VR in one field containing an area of recently cleared land, where vegetative indices responded to strong spatial differences, with S4 reaching a mean VR of 32.54% at 4 MZs in that field. Four management zones consistently provided the best balance between VR improvement and practical implementation, with five zones offering only marginal gains over four. Soil K showed the greatest zone differentiation across all fields, while soil pH was the most difficult soil property to differentiate due to its inherently low spatial variability. Overall, S2 was identified as the most broadly applicable and practically efficient delineation strategy, offering improved performance over EC alone while minimizing data requirements.

Artificial Intelligence (AI) Use Disclosure Statement

In preparing this thesis, the following Artificial Intelligence (AI) tool was used: Anthropic's Claude. This tool was used primarily for code syntax assistance and debugging, as well as grammatical and wording revisions. The author acknowledges full responsibility for the intellectual content of this work and has ensured that all AI-assisted sections have been reviewed and revised for accuracy and appropriate academic style. All AI-generated content was reviewed and validated for relevance, appropriateness, and accuracy before incorporation into the final document to maintain the scholarly integrity of this research.

Digital Accessibility Disclosure Statement

In the preparation of this thesis/dissertation, Microsoft Word Editor's accessibility tool was used to ensure this document complies with federal requirements. The author acknowledges full responsibility for the intellectual content of this work and has made a good faith effort to comply with digital accessibility requirements in publishing, wherein the nature of the content does not significantly change in order to do so. Furthermore, all content has been reviewed and revised to meet these requirements prior to final publication.

Acknowledgments

First of all, I would like to thank the people who have continuously encouraged me along my journey, my parents. My father, John Beasley, for showing me what it means to be the best in your profession, and my mother, April Beasley, for keeping me encouraged and committed to finishing the drill. To all of my close friends and family, I appreciate your help in keeping me going. Thank you to Greenpoint Ag, who were instrumental in the creation of this thesis by allowing us to work with growers and acquire the large amounts of data required for such a study.

Thank you to all my lab mates and graduate school friends who provided much needed assistance collecting hundreds of soil samples, including Jacob Sizemore, Shubhdeep Singh, Jarrett Hancock, Cole Byers, and Ravi Meena.

Thank you to Dr. Simerjeet Virk for his guidance and help on this journey, and for providing the opportunity to participate in numerous projects related to precision agriculture. The skills and experience I have gained under your guidance will stay with me for years to come. Thank you to my committee members, Drs. Audrey Gamble and Thorsten Knappenberger, for their expertise in the subject matter.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction and Literature Review

1.1 Introduction

Agricultural land in the southeastern United States is subject to significant spatial nutrient variability due to factors including soil type, elevation, and historical management practices. This variability often results in inefficiencies in crop input use, such as over- or under-application of fertilizers, excessive irrigation in poorly drained areas, or overuse of plant growth regulators (PGRs) in zones with lower vegetative vigor. These issues can further lead to increased field-level variability, lower yields, higher input costs, and greater environmental impact through nutrient leaching and surface runoff. Precision Agriculture (PA) seeks to address these problems by implementing Site-Specific Management (SSM) strategies that can address spatial variability within agricultural fields. According to the International Society of Precision Agriculture, PA is a management strategy that collects, processes, and analyzes spatial, temporal, and individual-level data, and integrates this with other sources of information for data-driven decisions that improve input efficiency, productivity, profitability, and sustainability (ISPA, 2022). PA relies on a variety of tools to support this approach, including the use of GPS technology, data collection from sensors and equipment, decision support systems, and variable-rate (VR) application technologies (Pedersen & Lind, 2017). The growing availability of farm data management software, also known as Farm Management Software (FMS), and other compatible hardware has allowed producers to incorporate these technologies into their routine management practices. This adoption has led to increased crop yields, reduced labor requirements, and overall greater farm productivity.

One of the more impactful PA practices is the ability to apply soil amendments, particularly fertilizers, at variable rates based on localized field conditions. VR applications of soil amendments and macronutrients such as phosphorus and potassium can be paired with management zones (MZs) within fields to target these inputs more effectively. These SSM approaches can reduce nutrient waste and address spatial variability more efficiently than traditional uniform methods, which involve the uniform application of lime and fertilizer across the field. As VR technology and spatial data analytics tools become more accessible, their use among growers and crop advisors across the United States is expanding (McFadden et al., 2023). In the southeastern US, where soil variability is often more pronounced, these technologies allow growers to be more selective in managing variability on their farms. The implementation of VR strategies enables farmers to reduce excess nutrient application without sacrificing yield performance. Consequently, there is a rising interest in adopting SSM strategies such as utilizing MZs for precision soil sampling and estimating in-field nutrient variability. However, one of the continuing challenges is the suitability and effectiveness of different spatial data layers for delineating MZs to inform accurate VR prescriptions for fertilizer and agricultural limestone.

MZs have existed within the agricultural community for well over 50 years, although the methodology for their delineation and the use of various data layers have evolved over time (Escadafal, 1993; A. Mallarino, 2001; Nawar et al., 2017; Schepers et al., 2004). In the past, a common method for creating management zones was to use the growers' historical knowledge of the variability within a field and then divide the field into two or three zones based on this information (Fleming et al., 2000). As spatial data have become increasingly abundant in precision agriculture, the use of soil and crop characteristics has become more common for the delineation of management zones. Several researchers have explored the creation of MZs using various spatial data layers such as historical yield data, pH, soil nutrient levels, soil type, and elevation (Coleman, 2021; Flowers et al., 2005; Mallarino & Wittry, 2004; Schepers et al., 2004; Tucker, 2024). While MZs can be based on only a single layer of spatial data, combining multiple relevant layers often provides a more robust representation of variability within the field. Although several studies have investigated MZ delineation based on individual layers, assessments of integrating multiple spatial layers remain limited, and even fewer have been conducted recently, especially in the last decade. Research in the Southeast is largely limited to single EC layers or to grower knowledge of nutrient MZs, with more research available on in-season irrigation and nitrogen management zones (Haghverdi et al., 2015; Taylor et al., 2007; Thompson et al., 2004; Tucker, 2024). Determining which spatial data layers or combinations are most effective for zone delineation remains an active area of investigation. In addition to selecting appropriate soil and crop properties, the accuracy of spatial data presents another major challenge. Crop yield and elevation data can often be downloaded from modern machinery, but data quality is frequently limited by poor calibration accuracy and GPS precision.

Precision soil sampling—grid- or zone-based—is often necessary to accurately determine nutrient variability within agricultural fields. Grid sampling is typically recommended at grid sizes of 0.4 to 1.01 ha (1-2.5 ac) to adequately capture spatial variability. However, this sampling intensity can be expensive and time-consuming, both in labor and in analysis costs. Precision soil sampling based on MZs offers a more economical alternative, provided the zones are homogeneous in soil properties. Maximizing homogeneity within zones requires careful selection and analysis of spatial layers correlated with nutrient availability, enabling accurate estimates of nutrient variability within fields. To better understand the potential of these MZ-based sampling strategies, research is needed on a regional scale. A study conducted across the southeastern US, using high-resolution spatial data and validated nutrient measurements, would provide insight into the feasibility and accuracy of MZs for precision soil sampling.

1.2 Site-Specific Nutrient Management

Site-Specific Nutrient Management (SSNM) is an integral part of precision agriculture and can be defined as an approach to applying nutrients to crops on an as-needed basis. When implementing SSNM, the application and management of nutrients are dynamically adjusted to meet crop needs based on location and growth (Dass et al., 2014). Additionally, the practical application of this practice requires machines capable of applying fertilizers at varying rates and, more importantly, the creation of accurate prescription maps that reflect the nutrient requirements within a field. Precision soil sampling is commonly used to determine the nutrient variability in agricultural fields, and numerous studies have investigated the effectiveness of various precision soil sampling strategies (Coleman, 2021; Flowers et al., 2005; Mallarino & Wittry, 2004; Shaner et al., 2008). Before the widespread adoption of precision soil sampling, most farmers collected composite field samples, which represent soil cores combined across the field to determine nutrient levels. This traditional soil sampling method resulted in a uniform application of nutrients across the field. Fields were sometimes divided into large areas based on historical knowledge, but these zones were not data-driven. Over the years, the increased use of GPS, in conjunction with soil sampling and sensing technologies, has enabled sampling to become more precise. Remote sensing capabilities in agriculture have also increased significantly, enabling the collection of significant amounts of soil and crop data (Sishodia et al., 2020). The increase in farm size in the US has also allowed farmers to adopt and expand the use of precision technology. A recent review by McFadden et al. (2023) reported that 50% of farmers in the third quintile by cultivated area have implemented precision agriculture methods, compared with less than 25% of smaller farmers. Additionally, the adoption of PA across farms has increased from 5.3% in 1996 to 44% in 2018. Coleman (2021) found that implementing PA methods decreased temporal yield variation by 30%. As a result, tillage and pesticide costs were also reduced. However, Mcfadden et al. (2023) also found that fertilizer costs associated with implementing some PA methods were increased, but these costs were associated with increased application of P and K fertilizers due to nutrient-deficient areas being revealed through PA. With the total costs of lime and fertilizer more than doubling since early 2021, farmers are expected to increase the use of variable-rate application from 37.2% in 2022 to 66.2% in 2024-2025 (Wang et al., 2023). The ability to collect large amounts of spatial data has enabled more farmers than ever to rely on precision strategies such as SSNM, making it essential to implement these practices effectively. A successful implementation of SSNM begins with reliable nutrient maps, which can be achieved through different precision soil sampling methodologies.

1.3 Grid-Based Soil Sampling

One of the most commonly employed precision soil sampling strategies in the southeastern United States is grid sampling. Grid soil sampling is a method in which a uniform grid is superimposed on a field (Figure 1.1) and soil samples are collected at the center of each grid cell, using GPS and Geographic Information Systems (GIS) for precise georeferencing. This approach remains prevalent in the Southeast due to its ease of implementation using widely accessible FMSs such as AccuField (Greenpoint Ag, 2021), Climate Fieldview (Climate Corporation, 2021), and AgLeader Spatial Management System (SMS) (Agleader Technology, 2004), which integrates well with most existing PA technology. Grid sampling has been extensively evaluated in the literature (Asare & Segarra, 2018; Flowers et al., 2005; A. P. Mallarino & Wittry, 2004; Virk et al., 2025), with studies confirming its ability to generate accurate nutrient distribution maps when properly executed. The effects of different parameters, such as sampling density, temporal frequency, and sampling location, have also been investigated to determine their impact on data quality and operational efficiency (Buscaglia & Varco, 2003; Thompson et al., 2004).

Figure 1.1. Example of soil sampling maps based on 2.0 ha (left) and 0.1 ha (right) grid sizes.

For effective grid sampling, grid size is particularly important, as it affects both the spatial resolution of soil nutrient data and the economic feasibility of sampling operations. Previous research has suggested that grid sizes around 0.4 ha (1 ac) are generally sufficient to capture meaningful spatial variability in soil phosphorus (P), potassium (K), and pH levels in the Midwestern United States (Flowers et al., 2005; Mallarino & Wittry, 2004). More recent studies have expanded on these findings, indicating that a grid size of 0.4 to 1.0 ha can be used depending on field conditions, crop requirements, and input costs (Virk et al., 2024). This study suggested that grid sizes above 0.4 ha can substantially reduce the accuracy of VR lime, P, and K, while maintaining similar overall costs. However, a study by Asare and Segarra (2018) found that around 50% of growers still use 1-ha grid sizes, due to lower sampling costs and labor requirements. While it was reported that most growers use 1-ha grids, the authors found no consensus on the optimal grid size. Consequently, selecting an appropriate grid resolution is critical not only for capturing relevant spatial variability in nutrients but also for ensuring efficient use of capital and labor within agricultural systems. More recent research in the Southeast has reported that a 1-ha grid size is optimal for efficient sampling while still maintaining spatial accuracy (Tucker, 2024). A study conducted by Coleman (2021) found that the most economically efficient grid size in the Coastal Plain region of South Carolina was approximately 0.4 ha at $304/ha for labor and fertilizer, though this could vary with soil characteristics. The authors reported that the traditional sampling method, which involves combining samples across a field, was by far the most expensive at $387/ha. Mallarino and Wittry (2004) found in a study over eight fields in the Midwestern US that smaller grid sizes directly revealed more variability. However, the authors also concluded that no single approach was effective across all eight fields that were evaluated.

In addition to grid size and economic considerations, soil sampling frequency and collection methods are also considered important factors. A general recommendation in the Midwestern United States is to soil sample once every 5 years, but this can vary significantly in the southeastern US based on the crop rotation (Thompson et al., 2004). The lower nutrient-holding capacity and high acidity of southeastern soils both contribute to the need for more frequent sampling. During grid sampling, samples can be collected using a cell or point method. Cell sampling involves collecting 10-15 cores randomly across a single grid area. These cores are then compiled into a single composite sample. In contrast, the point sampling method involves gathering 10-15 cores within a 15.2 m (50 ft) radius of the grid centroid and then compiling them into a single composite sample. Thompson et al. (2004) investigated both methods thoroughly and reported that point sampling can create potentially skewed results due to its small sampling area. Flowers et al. (2005) also investigated both methods and reported similar findings that grid cell sampling was more effective than the point method. While grid sampling has improved significantly over the years, cost reduction remains necessary due to the large number of soil samples required to accurately capture in-field nutrient variability. While labor costs have been increasing, overall application costs for phosphorus, potassium, and agricultural limestone have been reduced by using grid sampling strategies (Coleman, 2021). Soil sampling based on MZs has the potential to further reduce sampling costs, especially labor, while still maintaining an accurate representation of nutrient levels for effective VR fertilizer applications.

1.4 Zone-Based Soil Sampling

Soil sampling based on MZs is rapidly gaining popularity for determining soil nutrient levels in agricultural fields. MZ-based sampling relies on combining homogenous areas of the field by using one or more spatial data layers (Figure 1.2), such as soil electroconductivity (EC), soil survey data, topography, yield, aerial imagery, or farmer knowledge (Amer et al., 2021; Shaner et al., 2008; Taylor et al., 2007). Due to the increased adoption of PA technologies such as yield monitors and Unmanned Aerial Vehicles (UAVs), growers and crop advisors have access to more data than ever to create zones that reflect within-field variability. Compared to grid sampling, zone sampling allows for a reduction in the number of samples per field while maintaining similar levels of accuracy (Asare & Segarra, 2018). Research has shown that the increased availability of FMS has made access to and utilization of different spatial data layers related to field properties and machinery operations much easier over the past few years. This has made it significantly easier for growers and crop advisors to manage and implement zone-based management strategies for site-specific application of crop inputs, including fertilizer (Nawar et al., 2017).

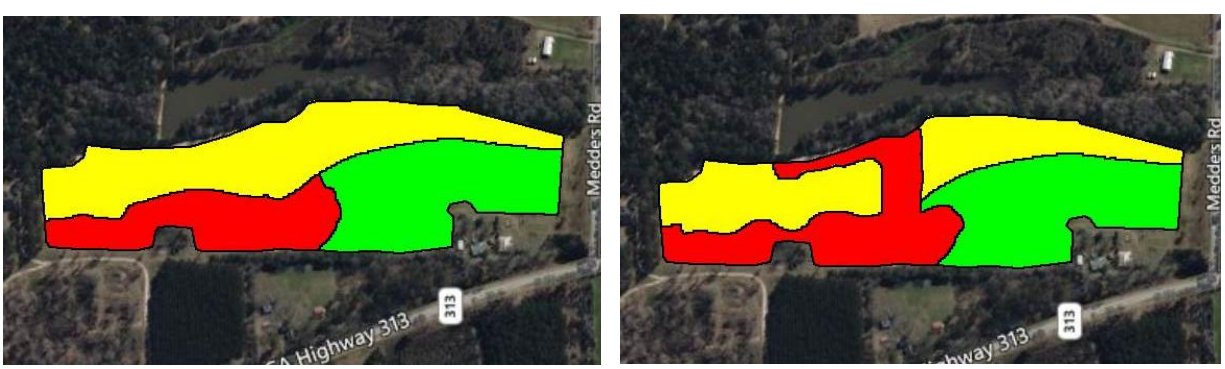

Figure 1.2. Maps depicting soil sampling zones delineated from soil EC and elevation (left) and soil EC, elevation, and NDVI imagery (right) in a field.

1.4.1 Soil Electrical Conductivity (EC)

One of the most widely available soil properties for creating MZs is soil electrical conductivity (EC) (Coleman, 2021; Schepers et al., 2004). Soil EC is correlated to soil texture, CEC, and moisture content (Neely et al., 2016). This can be beneficial in zone creation, as it helps identify areas with high nutrient and high water-holding capacity. Schepers et al. (2004) found that EC data can effectively map water content, salinity, and clay (%) in soil. One of the most common sensors used in soil profiling is the Veris 3100 (Veris Technologies, Division of Geoprobe Systems, Salina, KS) for soil EC collection (Setiawan et al., 2022). This method of EC data collection provides a high-resolution map by gathering EC across the field. This data is then interpolated using kriging or Inverse Distance Weighting (IDW) methods (Lu & Wong, 2008; Oliver & Webster, 1990) to create a raster data layer. When collecting EC data, most sensors distinguish between shallow and deep EC data. Shallow EC is defined as the EC measured within 30 cm of the soil profile, whereas deep EC extends to 90 cm (Peralta & Costa, 2013). Both can be used separately to create MZs; however, deep EC has a stronger correlation with yield than shallow EC in the Southeast (Coleman, 2021). In a study by Jaynes et al. (2005), apparent soil EC was used to predict a 5-year average of soybean yield. The authors found that soil EC accurately represented soybean yield across 80% of the evaluated area. Using a similar methodology, the authors also created yield maps based on measured soil EC for an adjacent field where no previous yield data had been collected. Peralta & Costa (2013) used deep soil EC to delineate management zones that reflected the variability in certain soil attributes. The authors evaluated two fields and created three MZs in each by using high, medium, and low EC levels. These delineated zones represented significant variation in soil pH and in some exchangeable cations, such as Ca^{2+}, Mg^{2+}, and Na^{+}. However, they also found that EC could not consistently identify P and K variability. One possible explanation is that the variability among P and K was limited in both fields, rendering MZs unsuitable for this purpose. While a large body of research uses EC data, only one of these studies (Coleman, 2021) was carried out in the Southeast. This lack of research in the region highlights the need for more comprehensive studies.

1.4.2 Crop Yield

MZ delineation based on crop yield is a common method for creating zones that reflect field productivity. As yield monitors have been around since 1993, nearly 50% of US growers have implemented yield monitoring technology on their harvest equipment (McFadden et al., 2023). Yield data is also often utilized when other types of data are unavailable. When using yield for MZ delineation, Flowers et al. (2005) recommended creating MZs tailored to the maximum production a particular area of the field can achieve. This is usually achieved by using 5-year normalized yield data, which eliminates the temporal variation that is common in yield datasets. Yield data can help reduce excessive input use when other issues, such as drainage or soil texture, are generally yield-limiting factors that prevent the crop from maximizing its potential. Boydell & McBratney (2006) found yield to be a valuable layer for creating accurate MZs, as it can indicate problem areas within the field. Sudduth et al. (2012) suggested that while yield data can be useful for MZ delineation, it is important to ensure the data is free of outliers and errors.

When creating MZs based on yield, Flowers et al. (2005) found lower coefficient of variation (CV) values in year-to-year nutrient levels by 50-77% in cotton compared to traditional whole field management, where the nutrient values were measured through a control of randomly sampled grid points throughout the field. The CV across several years of yield data was calculated to assess temporal yield variability in this study, and similar CVs were grouped to create MZs. The authors demonstrated its effectiveness in creating distinct zones that accurately reflected variability in phosphorus, potassium, and pH across several fields. When analyzing the economic efficiency of MZs based on yield, the authors reported that profitability was highly dependent on the crop, with higher-value crops like corn yielding higher profits than wheat or barley. Hornung et al. (2006) evaluated several methods of MZ delineation using yield data. Management zones across three sites were created by dividing each field into areas of high, medium, and low average yields based on yield data from the previous year. The authors utilized k-means clustering for the zone delineation. Interestingly, the authors found that the performance of MZs decreased when combined with other data layers. The authors indicated that the weighting of different layers may have caused these issues, as some layers may be more important for predicting crop yield. One potential issue arising from this study is the lack of temporal stability in the yield data, as only a single year of data was used for MZ delineation. While a number of studies have evaluated yield-based MZs, little work has been conducted in the Coastal Plains region of the Southeast, creating an opportunity for further research.

1.4.3 Soil Survey Data

Soil series is another data layer commonly used in MZ delineation, as it has been shown to significantly affect nutrient-holding capacity, drainage, and overall field productivity (Rojo Baio et al., 2019). Soil series can be described as soils that have horizons that are similar in composition, thickness, and arrangement (Ditzler, 2017). Most research evaluating soil series uses publicly available SSURGO maps for MZ delineation, provided by the National Cooperative Soil Survey (Soil Survey Staff, 2025). A review by Nawar et al. (2017) showed that, while soil survey maps are low-resolution when used for row-crop applications, they were still able to create effective MZs for larger field sizes. The authors delineated zones based on soil series for in-season nitrogen management and found that using MZs reduced nitrogen use by 13% without a significant reduction in yield. Additionally, N-use efficiency increased from 0.54:1 to 0.45:1 kg/N per bushel. However, the authors cautioned that, when comparing MZs from soil series to traditional directed grid sampling, the zones were ineffective at representing most field-level nutrient variability. Another study by Escadafal (1993) shows that soil color can also be used for soil series classification, and therefore for MZ delineation. It was also reported that soil color could be effectively sensed using satellite or aerial imagery, increasing its viability as a dataset for MZ delineation.

Hornung et al. (2006) compared MZs delineated using historical yield data with MZs derived from soil color and reported that the latter were more accurate at capturing the temporal variability in yield. In addition to temporal stability, the authors found that field productivity, in relation to yield, was better classified using soil color than historical yield data. Farmaha et al. (2020) suggested that soil survey maps can be used for zone sampling applications, and also recommended incorporating additional historical data to refine MZs. Gozdowski et al. (2014) evaluated variability in measured sand and silt content against MZs delineated from yield and found significant differences in both sand and silt among the MZs. This study suggests that sand and silt content were viable datasets for MZ delineation. Overall, MZs based on soil series show a mixed ability to create effective MZs and should often be combined with other layers for maximum effectiveness.

1.4.4 Elevation

Elevation is another widely available spatial dataset often used to delineate MZs. Using elevation for MZ delineation can be especially useful for growers with limited technology, as this data can be easily gathered online through open-source applications. A study by Castrignanò et al. (2011) used elevation to predict soil pH over a 1.5-ha area. The authors found that the digital elevation model (DEM) partially explained the variance in soil pH within the area. For soil nutrient values, a study by Baxter and Oliver (2005) evaluated the ability of elevation to predict soil inorganic N. The study conducted over two fields used three types of kriging weightings to compare the datasets. The authors found that elevation data partially explained the variance in soil inorganic N. Guo et al. (2012) showed strong correlations between elevation and slope and yield using 5 years of temporal data. Elevation can be an advantageous data layer for MZ delineation due to its ability to represent field variability and its ease of acquisition, but it is often better to use it in conjunction with other spatial data for accurate depiction of in-field variability.

1.4.5 Aerial Imagery

Aerial imagery offers several advantages for collecting spatial data. Many data layers have become viable options for MZ delineation thanks to the recent increase in drone availability. Aerial Imagery has been found to be effective for collecting many data layers, including NDVI imagery for crop health and vigor, soil color, LiDAR for crop height, plant density, and several other crop properties (Weiss et al., 2020). Additionally, some companies offer commercially sensed data with high temporal resolution for in-season crop monitoring. A study conducted by Song et al. (2009) evaluated NIR and visible imagery (2.4 m resolution) and panchromatic bands (60-cm resolution) and found that aerial imagery was able to reflect not only the growth of wheat within the field, but also the spatial variation in soil properties, including organic matter, available phosphorus, and extractable potassium. Fleming et al. (2000) utilized visible-spectrum satellite imagery to manually delineate zones of high, medium, and low productivity areas within a single field. However, significant inaccuracies were identified when zones were defined manually, as higher productivity areas were grouped together with medium or low productivity areas. Nawar et al. (2017) evaluated UAV-sensed NDVI imagery (crop coverage) using the leaf area index to create management zones.

However, the study concluded that this method was only somewhat useful in variable-rate applications, specifically for nitrogen inputs. A review of the available research has shown that aerially collected imagery can be highly variable in its ability to generate effective MZs and should be used alongside more stable data layers, such as soil EC, when possible.

1.4.6 Delineation Methods for Management Zones

The different types of spatial data that are openly available or easily collected often come in various formats and can sometimes exhibit large skews due to issues during data collection, such as excessive cloud cover, harvester speed, or image resolution (Sishodia et al., 2020; Taylor et al., 2007). Therefore, data cleaning is critical to ensure data quality and reliability. Limiting aerially sensed data to clear days and removing outliers from yield data can help create management zones that accurately represent variability within the field. To delineate accurate MZs, normalization of data layers is also emphasized to ensure consistency across datasets and minimize biases arising from differences in scaling or units. A study by Buscaglia & Varco (2003) reported that log transformation of variables can be used for data normalization if variables are not normally distributed. Several statistical methods have been used to delineate management zones, employing different data layers and weighting approaches. Nawar et al. (2017) analyzed several clustering methods and found that fuzzy c-means was a useful approach for delineating zones using soil fertility data from lab analysis. Additionally, the authors stated that k-means clustering was also effective when classifying zones using historical yield data. The study also found that smoothed fuzzy classification was effective for creating homogeneous zones with minimal variance within the delineated area. In general, there is no consensus on which statistical method to use for delineating MZs; however, Javadi et al. (2022) found that k-means clustering performed well in this context. Gallardo-Romero et al. (2023) evaluated four supervised clustering algorithms for zone delineation, including random forest (RF), classification and regression trees (CART), gradient-boosted trees, and support vector machines (SVMs). The study also evaluated k-means as the single unsupervised algorithm. All methods were used to create MZs from various aerially sensed imagery, such as NDVI and NDRE. The authors found that the RF and CART models exhibited the highest level of accuracy, while the SVM model was the least effective.

1.5 Rationale

A review of the existing literature showed substantial research on grid soil sampling, including its effectiveness and economic efficiency. Grid sizes ranging from 0.4 to 4.0 ha (1-10 ac) have been evaluated by several researchers, and recommendations (≤1.0 ha) for achieving a good balance between economic efficiency and spatial accuracy have been provided. For zone soil sampling, however, several questions remain unanswered, especially in the Southeastern region. While numerous studies have investigated zone delineation, they have identified several common issues, including the selection of appropriate spatial data layers for zone delineation. While a vast array of spatial data is available today, Gozdowksi et al. (2014) recommended reducing multivariate analysis to two or three layers for ease of implementation and computation. This is practical for growers as it allows zone implementation when limited data is available. Amongst different spatial layers, many have been identified as viable datasets for zone delineation. However, some datasets are more relevant to nutrient management than others, and their relevance varies by geographical region. Additionally, many studies have reported differing zones using the same soil and crop properties due to differences in the delineation methodology (Duffera et al., 2007; Javadi et al., 2022; Nawar et al., 2017). While some exploration has been conducted into MZ delineation for zone-based soil sampling, the results have varied significantly across regions, soil types, and spatial datasets used in delineation. Also, most research in this area has been conducted in the Midwestern United States or outside the country entirely, with minimal research undertaken within the Southeast, particularly in the Coastal Plain region. The proposed study aims to address some of the existing research gaps in zone soil sampling in this region. The study aims to explore the potential of various spatial soil and crop properties for MZ delineation and to validate the MZs for estimating in-field nutrient variability. This study will help improve the understanding of the feasibility of management zones for soil sampling and their potential to improve nutrient management in the southeastern US.

1.6 Objectives

The main goal of this research was to investigate the potential of various crop and soil spatial properties to delineate management zones for precision soil sampling and estimate nutrient variability within agricultural fields in the southeastern US. The specific objectives of this research were as follows:

1) To correlate different soil and crop spatial properties with within-field nutrient levels and determine their feasibility for delineating management zones
2) To validate the effectiveness of management zones delineated from various soil and crop spatial properties in predicting in-field nutrient variability in agricultural fields

Chapter Two

Evaluation of Different Soil and Crop Spatial Properties for Predicting In-Field Nutrient Variability and their potential for Management Zones for Precision Soil Sampling in the Southeast

2.1 Abstract

Precision soil sampling plays a critical role in site-specific nutrient management in precision agriculture. While grid sampling has become well established across the Southeast for this purpose, it is often expensive and labor-intensive. Zone sampling has recently gained interest among growers for its potential to reduce labor costs while maintaining accurate spatial representation of soil nutrients. To investigate the viability of zone sampling as a substitute for grid soil sampling, a study was conducted across nine fields (ranging from 9.6 to 37.8 ha) in the Coastal Plain region of Southern Alabama and Georgia. High-density grid soil samples (0.1 ha) were collected across all fields to establish baseline nutrient variability, including soil pH, phosphorus (P), and potassium (K). Spatial data layers evaluated included shallow and deep soil electrical conductivity (EC), elevation, normalized difference vegetation index (NDVI), normalized yield, and soil series. All datasets were standardized to a 32 × 32 m resolution to match the soil sampling grid. Pearson correlation analysis and Random Forest (RF) modelling were performed to assess relationships between the spatial layers and soil nutrient levels on a field-by-field basis. Soil P and K exhibited considerable spatial variability across all nine fields, with CV values reaching up to 46% and 39%, respectively, while soil pH showed comparatively little variability, with CV values not exceeding 6.8%. Pearson correlation results indicated that no single spatial layer showed consistent relationships across all fields. Deep EC demonstrated significant correlations with P, K, and soil pH in seven or more fields, and elevation showed significant associations across five to seven

fields, depending on the soil property, while soil series, NDVI, and yield showed more limited and field-specific performance. RF modelling indicated R^2 values between 0.04 and 0.78, with soil K models showing the most consistent metrics and soil pH models performing the weakest, reflecting the low inherent spatial variability of pH across the study fields. Across both analyses, deep EC and elevation were the most frequently identified predictors of nutrient variability, and a combination of these two layers is recommended as the most practical and broadly applicable basis for management zone delineation in Coastal Plain soils. However, the absence of a universally predictive layer set across all fields suggests that zone delineation strategies should be evaluated on a field-by-field basis rather than applied uniformly across a farm or region.

2.2 Introduction

The high spatial variability across agricultural farmland in the Southeastern United States poses significant challenges for growers, as it often leads to reduced yield and input efficiency. This variability can be associated with soil properties such as moisture, nutrients, pH, soil texture, and nutrient-holding capacity, and is often significant regardless of field size (Beckett & Webster, 1971). Within-field variability can be exacerbated by continuous uniform application of crop inputs over a long time, often resulting in over- or under-applications in certain areas of the field (Sawyer, 1994). To address this in-field nutrient variability, particularly with costly inputs like agricultural limestone, phosphorus, and muriate potash, variable rate (VR) applications are commonly used for site-specific management of nutrients in agricultural fields (Pedersen & Lind, 2017). The VR application enables fertilizer to be applied at the right rate and in the right place in the field, thereby facilitating the effective implementation of a 4R nutrient management strategy (Fixen, 2020). Proper implementation of VR applications is highly dependent on accurately depicting nutrient variability within the field, which in turn relies on precision soil sampling strategies.

Precision soil sampling can be broadly classified into two main types: grid and zone sampling (Ackerson, 2018). Currently, grid soil sampling is the most commonly employed precision soil sampling strategy in the southeastern United States (Walton et al., 2008). In grid soil sampling, a field is divided into uniform-sized grids, and soil samples are collected at the center of each grid cell (Flowers et al., 2005) with the aid of GPS-enabled systems for georeferencing. This approach remains prevalent in the Southeast due to its ease of implementation using widely accessible spatial data management software such as AccuField (Greenpoint Ag, 2021), Climate Fieldview (Climate Corporation, 2021), and AgLeader Spatial Management System (SMS) (Agleader, 2004), which integrate well with most existing precision agriculture technology and systems. Given its widespread adoption, grid sampling has been extensively evaluated in the literature, with studies confirming its ability to generate accurate nutrient maps when properly executed. Various factors related to grid sampling, such as sampling density, temporal frequency, and sampling location, have been studied to determine their impact on data quality and operational efficiency (Flowers et al., 2005; Mallarino & Wittry, 2004; Virk et al., 2024).

In contrast, zone sampling involves dividing the field into smaller, homogenous areas that can be managed uniformly and collecting composite soil samples within each zone (Shaner et al., 2008). Zone soil sampling can provide the benefits of high-resolution grid sampling, while reducing the number of in-field samples and consequently sampling costs (Valente et al., 2024). Precision soil sampling based on management zones (MZs) offers a more economical alternative, provided the zones are homogeneous in soil properties. Maximizing homogeneity within zones requires careful selection and analysis of spatial layers correlated with nutrient availability, as different layers may have varying relationships with soil nutrients (Pusch et al., 2023). Management zones for soil sampling can be based on various spatial data sources, including crop yield, soil pH, EC, elevation, and soil type (Coleman, 2021; Flowers et al., 2005; Mallarino & Wittry, 2004; Schepers et al., 2004). In recent years, the rise of remote and in-field sensing technologies has provided access to large amounts of spatial data related to soil and crop properties (Nawar et al., 2017). Most studies on zone sampling have shown that specific layers are effective only within a single field and cannot be generalized to larger areas or the whole farm. Additionally, several researchers focused solely on using a single layer for MZ delineation, whereas multiple spatial layers can often yield more refined zones (Gozdowski et al., 2014). Proper implementation of zone sampling requires adequate delineation of MZs, making this soil sampling strategy less widely adopted than grid sampling. For this reason, studies investigating the potential of zone sampling, especially in the southeastern US, have also been limited, with only a few studies evaluating fields in this region (Coleman, 2021).

Past studies have identified several common issues in the MZ delineation process, the most significant of which is the selection of appropriate spatial data layers for zone delineation. While a vast array of spatial data is available today, Gozdowski et al. (2014) recommended reducing multivariate analysis to two or three layers to facilitate implementation and computation. This is also more practical for producers as it allows zone implementation when limited spatial data is available. Among the various spatial layers, many have been identified as viable datasets for zone delineation (Gozdowski et al., 2014; Nawar et al., 2017; Shaner et al., 2008). However, some datasets are more relevant to nutrient management than others, and their relevance also varies by geographical region. Additionally, many studies have reported differences in MZ’s using the same soil and crop properties due to differences in the delineation methodology (Duffera et al., 2007; Javadi et al., 2022; Nawar et al., 2017). Many previous studies have examined only one or two fields; however, to better understand the potential of these MZ-based soil sampling strategies, research is needed on a regional scale. A study conducted across multiple fields, using high-resolution spatial data and validated nutrient measurements, would provide insight into the feasibility and accuracy of MZs for precision soil sampling. Therefore, this study was undertaken to investigate the potential of different spatial layers for delineating management zones that can provide an accurate depiction of soil pH, P, and K. Given the high spatial variability of Coastal Plain soils, this study included several fields that represented the prevalent soil types and soil variability prevalent within the fields in this region. The main objective of this study was to determine the correlation between different soil and crop spatial properties and soil pH, P, and K variability within the fields, with the goal of identifying the spatial layers best suited for delineating MZs for precision soil sampling.

2.3 Materials and Methods

2.3.1 Experimental Sites

The data for this study were gathered across nine fields in the southeastern US in 2024 and 2025. The selected fields ranged in size from 9.6 to 37.8 ha (Table 2.1) and were located in the Coastal Plain region of the southeastern United States (three in Alabama and six in Georgia). The detailed information on the prevalent soil types in these fields is presented in Figure 2.1. These fields were selected randomly based on the criteria that they were representative of the soil types prevalent in the Coastal Plain region, with some known pH, P and/or K variability according to the producers knowledge. Most of these fields were in a cotton-peanut rotation in the previous years, where the field conditions were strip-till for cotton and conventional tillage for peanut. Five out of the nine fields were irrigated, and the four remaining were dryland.

Table 2.1 Location, size, and soil information on all nine fields used in this study.

Field	Longitude	Latitude	Size (ha)	Water Management
1	-83.4944	32.16946	30.1	Irrigated
2	-83.9045	32.03338	17.2	Dryland
3	-83.902	32.06156	14.9	Irrigated
4	-83.8647	31.65650	9.6	Dryland
5	-85.1683	32.02541	37.8	Irrigated
6	-83.9015	31.69454	11.8	Irrigated
7	-83.4668	32.18166	11.9	Dryland
8	-85.0894	32.05022	19.7	Dryland
9	-85.0676	32.05808	17.7	Irrigated

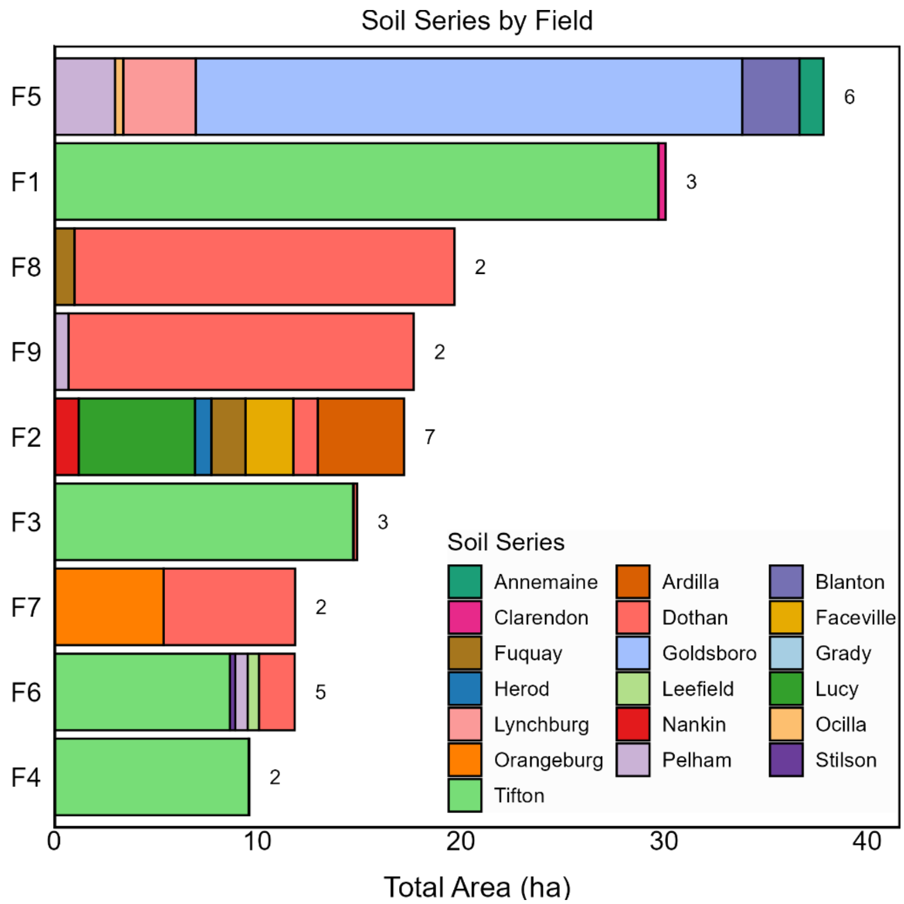

Figure 2.1. Soil type information for each field. The number in front of each bar indicates the number of soil types present within each field.

2.3.2 Spatial Data Acquisition

Soil and crop spatial data, including soil EC, elevation, soil type, NDVI, and yield, were collected for the selected fields. Soil electroconductivity (EC) data were collected across eight of the nine selected fields. Soil EC was collected using a Veris Technologies Mobile Sensor Platform (MSP) 3100 (Salina, KS) pulled with a UTV (Kubota RTX 900, Kubota Tractor Corporation, Grapevine, TX 76051). This UTV was equipped with a Trimble GPS 7500 auto-guidance system (Trimble, Westminster, CO) with Real Time Kinematics (RTK) capabilities, which enabled the georeferencing of spatial data with an accuracy of ±1 cm. Both shallow (30 cm) and deep (90 cm) soil EC data were collected in each field. In this instance, the sensor was set to collect one data point every second (1 Hz), and data was collected by making consecutive passes spaced 12 m apart within each field. The raw EC data were processed using Field Fusion software from Veris Technologies (Salina, KS) to remove any erroneous data and other outliers. In addition to shallow and deep EC data, the sensor also recorded elevation across the field at the same spatial resolution. Due to the high accuracy of the GPS/GNSS unit utilized, elevation data collected during this process were used as one of the spatial datasets for each field. For field 2 where soil EC data could not be collected, elevation was obtained from a John Deere CP 770 cotton picker (John Deere, Moline, IA) equipped with a high-accuracy RTK GPS unit.

Yield data was available for five fields in this study. Yield data was collected through the CP 770 cotton picker using an onboard yield monitor, and raw data was uploaded directly into the Accufield software (GreenPoint Ag, Decatur, AL). Only fields that had two or more years of available yield data were selected to create a stable normalized yield dataset. When cleaning yield data, the procedure outlined in Sudduth et al. (2012) was followed to normalize data used in the analysis.

Normalized difference vegetation index (NDVI) data is widely available through different Farm Management Information Systems (FMIS). In this instance, NDVI was accessed using the AgLeader SMS Advanced software (Ag Leader Technology, Ames, IA), which utilizes Sentinel-2 satellite imagery at a spatial resolution of 10 × 10 m to generate this crop index. The NDVI data was downloaded in August 2024, as most fields were at or near peak growth and canopy coverage during this time. The maximum allowed cloud cover was set at 10%; however, the maximum cloud cover downloaded for any field never exceeded 2%. Any non-vegetative or cirrus cloud cover pixels were removed, but no more than 10% of the total pixels were removed in any field.

Soil type data was acquired in a format similar to the NDVI data. Vector maps of soil types were downloaded through SMS Advanced software (AgLeader Technology, Ames, IA) using the soil survey data available from the Web Soil Survey (Soil Survey Staff, 2025). The downloaded data was clipped to each field boundary for further analysis. Table 2.2 provides information on various types of data collected and analyzed for each field.

Table 2.2. Information on different soil and crop specific spatial properties gathered for each field.

Field	1	2	3	4	5	6	7	8	9
Variable (Attribute)									
EC Shallow	X	-	X	-	X	-	X	X	X
EC Deep	X	-	X	X	X	X	X	X	X
Elevation	X	X	X	X	X	X	X	X	X
Normalized Yield	X	X	X	-	-	X	X	-	-
NDVI	X	X	X	X	X	X	X	X	X
Soil Type	X	X	X	X	-	X	X	X	X

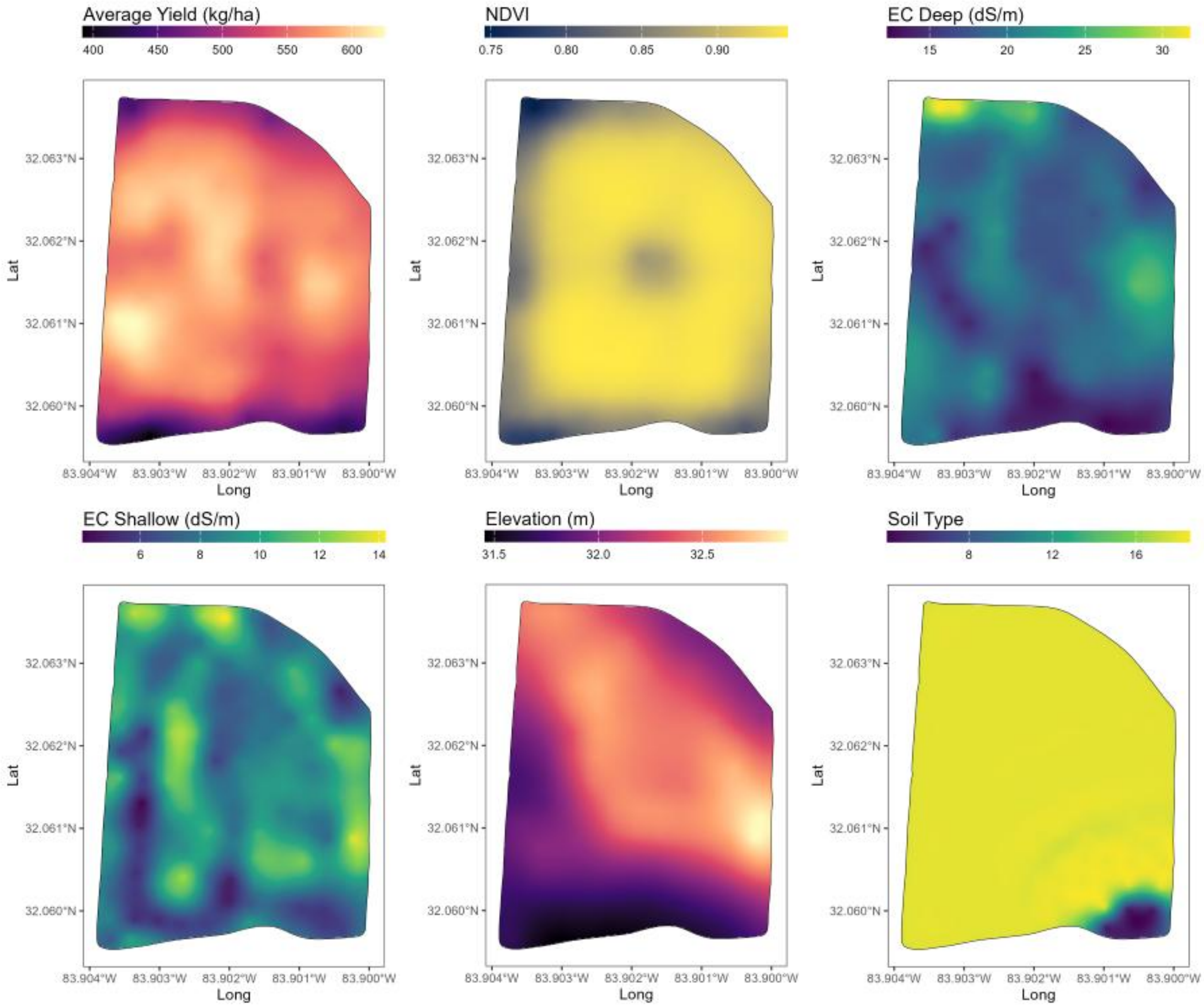

Figure 2.2. Interpolated maps of different soil and crop properties for Field 3.

2.3.3 Nutrient Variability - Grid Soil Sampling

To determine the existing nutrient variability (soil pH, P, and K) within each field, soil samples were collected using 0.1-ha grids. For each field, the field boundary was imported into the Accufield software (Greenpoint Ag, 2021), and 32 × 32 m (0.1 ha) grids were overlaid over these boundaries to create the soil sampling maps. These maps were then imported into the FieldAlytics Mobile app on an Apple iPad and connected to a BadElf GPS Pro unit (Bad Elf LLC, West Hartford, CT) for georeferencing samples. These GPS units have a 2.5-meter horizontal accuracy, which was sufficient for field navigation. The soil samples were collected from each field between January and March 2025, which is a typical sampling period in the region. Out of the nine fields, three fields were soil sampled following the peanut crop, and the remaining fields were sampled following the cotton crop. The soil samples were collected using the point sampling method, with 12 to 15 cores taken within each grid. All soil cores were taken at a depth of 15 cm and were mixed within each grid to form a composite sample, which was then placed in pre-labeled soil bags and sent to the University of Georgia's Agricultural and Environmental Services Laboratories (AESL) in Athens, GA. They were then subjected to Mehlich-1 extraction for soil macronutrients, including extractable phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), manganese (Mn), and zinc (Zn). The soil pH analysis was conducted using a 0.01 M $CaCl_2$ suspension, which provide results with higher stability compared to water dilution methods (Kissel & Sonon, 2011). Since this study focused primarily on correlating soil pH, P, and K with different spatial soil and crop properties, only values for these nutrients were retained for data analysis and soil mapping.

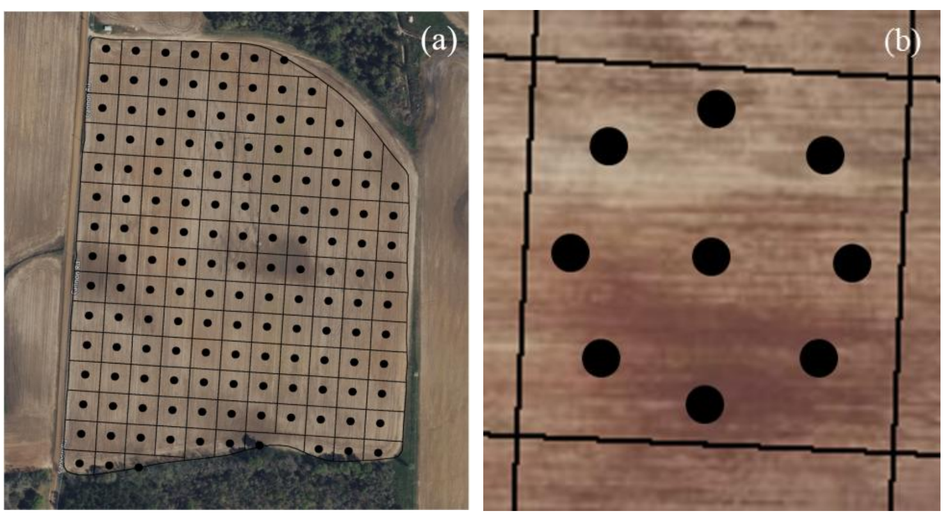

Figure 2.3. Soil sampling map (a) based on 0.1-ha grids for field 3; (b) an illustration representing where cores were taken in each grid. Each black dot represents a single core.

2.3.4 Mapping and Data Analysis

Various soil and crop datasets collected in this study were available at different spatial resolutions and formats. While some data had a spatial resolution of up to 10 × 10 m, such as NDVI imagery, the soil sampling data were collected at a high spatial density of 32 × 32 m. Additionally, the post-processed EC data were returned as point data across 12 m swaths. Therefore, all data sets were imported into SMS Advanced Software (Version 24.2, 2004) and interpolated using the Inverse Distance Weighting (IDW) method. The IDW interpolation method uses an algorithm that assigns data points higher weights depending on their distance from the interpolated area (Lu & Wong, 2008). After interpolation, the data layers were rasterized by resampling to a 32 × 32 m resolution to align with the soil sample data. This helped create a single data point per grid cell for each layer, matching the soil sampling data resolution. This method was applied to all raster datasets except the soil type dataset, which was available in a vector format. Therefore, additional processing was performed on the soil series data to create a usable dataset. The previously created 0.1-ha grids were exported as a shape file along with the soil survey data. This data was then imported into ArcGIS Pro software (ESRI, 2025), and a spatial join was performed between the soil type and the 0.1-ha grid layer. The largest area overlap of each soil type with the grid points was fully applied to each raster box. Then, the soil types were assigned numerical values, with each number representing the same soil type across all fields. This sometimes resulted in the elimination of some soil types as they were only present in a small, sliver polygon in mapping. Once all the data were resampled and represented at a 32 × 32 m resolution, they were exported as a CSV file for each field for further analysis. Figure 2.4 illustrates the workflow, showing various conversion and resampling techniques applied to the dataset across all fields.

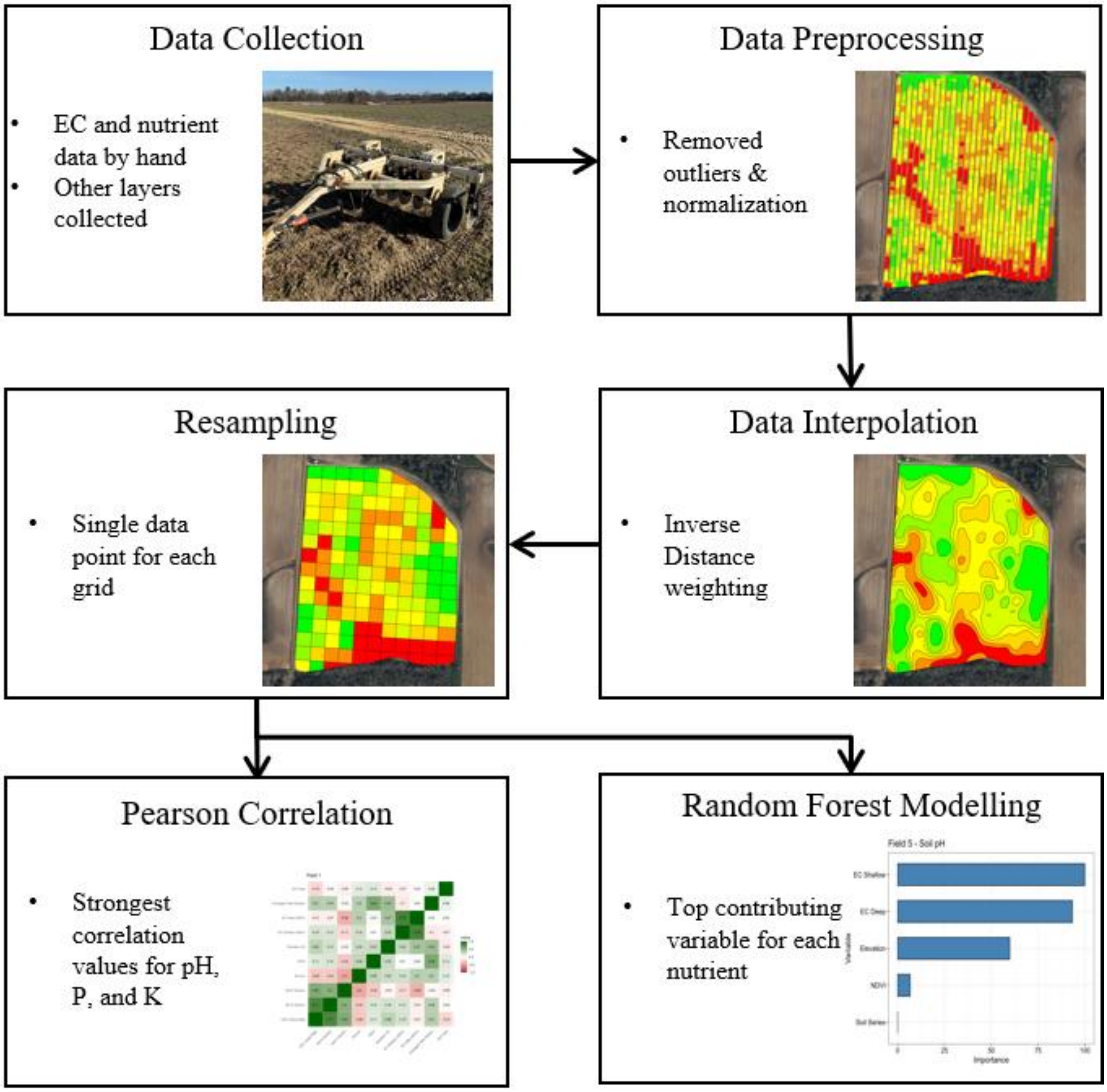

Figure 2.4. Flowchart showing the data-processing workflow used to analyze data for each field.

The CSV files containing data for all nine fields were imported into R Studio (RStudio 2025.05.1) using the *readr* package (Wickham et al., 2024). During initial data exploration, the mean, median, coefficient of variation, skewness, and kurtosis were calculated in R using the *e1071* package (Meyer et al., 2025) for pH, P, and K. For further analysis, the spatial property data (EC, elevation, yield, NDVI, soil series) were log-transformed to address skewness due to the variety of data scales. The transformed data were subjected to a Pearson Correlation analysis using the base R function *cor*, and plotted using the *corrplot* package. This analysis helped to initially identify any strongly correlated datasets with soil pH, P and K. To determine the most relevant data layers for management zone (MZ) delineation, a random forest model was developed for each nutrient in each field using the *ranger* package (Wright & Ziegler, 2017). For this modelling, data was partitioned into an 80/20 train-test split. The spatial data layers that showed a moderate to high correlation with nutrient levels were identified and discussed as having the greatest potential for delineating MZ's.

2.4 Results and Discussion

2.4.1 Within-Field Nutrient Variability

To understand the variability in soil pH, P, and K levels within each field, descriptive statistics including mean, median, range, coefficient of variation (CV%), skewness, and kurtosis were calculated (Tables 2.3, 2.4, and 2.5, respectively). The kurtosis and skewness values demonstrate the normality of the data (Dangal et al., 2019) where higher kurtosis values indicate a higher occurrence of outliers. Similarly, high skewness values indicate a longer tail, with negative values suggesting a leftward skew and positive values rightward skew in the data.

Soil pH showed the least variability among the three variables of interest, ranging between 4.9 and 7.6 (Table 2.3). Compared with the Auburn University soil fertility recommendations, only three fields (Field 4, 8 and 9) had an average soil pH level low enough (≤5.8) to warrant application of agricultural lime (Mitchell & Huluka, 2012). Fields 2, 6, and 9 exhibited the highest variability in soil pH, with CV values exceeding 6%. In terms of data distribution, fields 1 and 8 have higher skewness and kurtosis for soil pH, raising concerns about the assumption of normality. Median soil pH values were nearly identical (±0.1) to mean soil pH values across all fields, suggesting minimal outliers in the data. Figure 2.5 shows soil pH variability across the fields, and areas within orange or red have soil pH values low enough to require a lime application. Classifications for soil pH were as follows: very low (4.5-5.0), low (5.0-5.5), medium (5.5-6.0), high (6.0-6.5), and very high (6.5+). Fields 9 and 8 have substantially lower pH levels, with >50% of the field area requiring lime application. Fields 2, 4, 5 and 6 require a moderate amount of soil amendment, with areas requiring application ranging from 15 - 27%. The remaining fields had sufficient pH levels, with less than 12% of field areas needing soil amendments. Overall, soil pH variability was considerably low and had similar distributions across all nine experimental sites.

Table 2.3. Descriptive Statistics for soil pH for the fields used in this study.

Field	Range	Mean	Median	CV (%)	Skewness	Kurtosis
1	5.2 - 7.1	6.6	6.6	4.2	-1.25	2.78
2	4.9 - 7.1	6.3	6.4	6.5	-0.94	0.80
3	5.8 - 7.3	6.4	6.4	4.2	0.22	0.03
4	5.4 - 6.7	6.1	6.2	5.3	-0.42	-0.57
5	5.4 - 7.6	6.3	6.3	4.7	0.01	0.62
6	4.9 - 7.2	6.4	6.5	6.8	-0.57	-0.15
7	5.7 - 6.9	6.4	6.4	4.3	-0.24	-0.70
8	5.6 - 7.2	6.0	6.0	4.2	1.24	3.31
9	4.9 - 6.7	5.8	5.7	6.4	0.27	-0.51

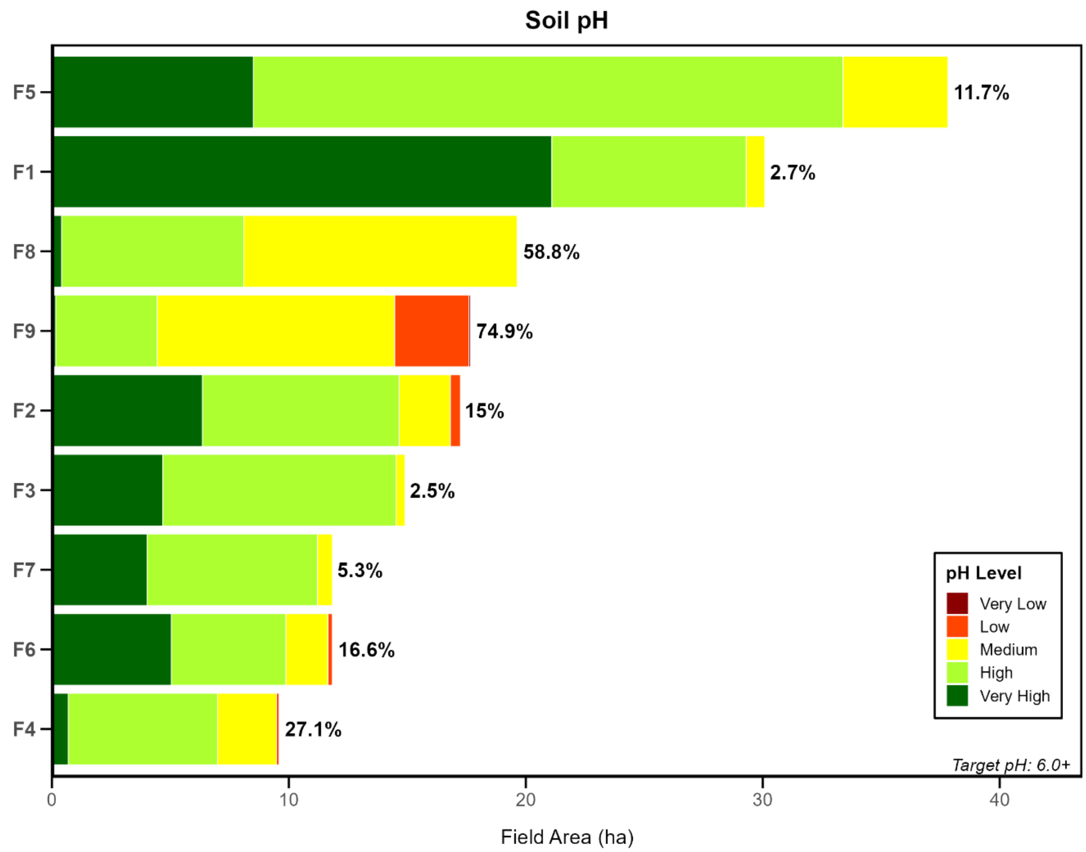

Figure 2.5. Area within each field classified into different soil pH levels. Percentage values in front of the bars represent the area within the field that requires an agricultural lime application.

Soil P showed considerably greater variability than soil pH, with values ranging from 7 to 524 kg/ha across all fields. Mean soil P values ranged from as low as 41 kg/ha in Field 7 to as high as 265 kg/ha in Field 5. The CV values ranged from 22% in Field 3 to 46% in Field 2, indicating substantial spatial variability among the fields. This high variability could partially be attributed to differences in soil texture and soil mineralogy, which influence P sorption capacity and availability (Veloso et al., 2023). Highly weathered, Fine-textured soils with higher clay content typically exhibit greater P fixation, reducing plant-available P, while sandy soils retain less P but maintain higher availability. Previous studies also partially attributed P variability to differences in soil texture (Marcaida et al., 2025). Median P values were similar to mean values across most fields, suggesting relatively symmetric data distributions. However, Field 1 exhibited higher skewness (1.19) due to a single observation exceeding the others by more than 100 kg/ha, indicating an outlier. Kurtosis values ranged from -0.32 to 5.47, with Fields 1 and 8 exhibiting the most pronounced peaks.

Based on Auburn University soil test recommendations (Mitchell & Huluka, 2012), only Field 6 and 7 had mean soil P concentrations low enough (<50 kg P_2O_5/ha) to warrant significant P applications, with well over 50% of the field area needing fertilizer. Figure 2.6 shows the classification of P values in each field, which were as follows: very low (0-28 kg/ha), low (28-56 kg/ha), medium (56-84 kg/ha), high (84-112 kg/ha), and very high (112+ kg/ha). Fields 1 and 4 had a moderate fertilizer requirement, with 15-25% of the field area requiring P applications. All other fields had much higher P levels throughout, with only 5% or less of the field area requiring fertilizer application. The soil P levels varied widely from field to field and did not exhibit patterns of variability similar to those of pH levels.

Table 2.4. Descriptive Statistics for soil P (kg/ha) for the fields used in this study.

Field	Range	Mean	Median	CV (%)	Skewness	Kurtosis
1	29-374	98	94	43	1.19	5.47
2	38-227	140	151	30	-0.52	-0.32
3	40-254	128	127	35	0.41	-0.23
4	24-138	72	67	33	0.56	-0.23
5	55-524	265	261	29	0.34	0.53
6	7-90	41	43	44	0.40	-0.10
7	9-122	50	48	46	0.77	0.40
8	81-332	153	146	22	1.45	3.90
9	8-188	94	93	29	-0.05	1.22

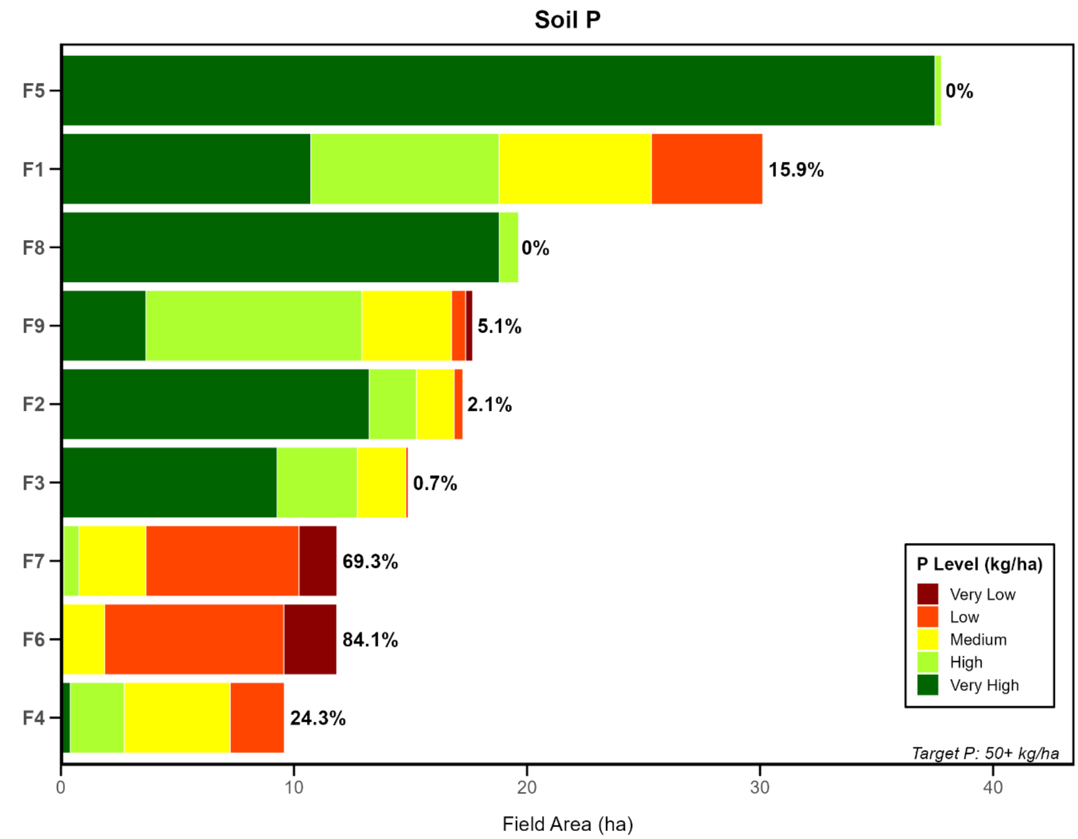

Figure 2.6. Area within each field classified into different soil P values. Percentage values represent the total amount of the field that requires fertilizer.

Similar to soil P, soil K showed considerable variability across the fields, with values ranging from 39 to 486 kg/ha. The mean soil K ranged from 127 kg/ha in Field 4 to 219 kg/ha in Field 3. The CV values ranged from 19% in Field 3 to 39% in Field 2, indicating moderate to high spatial variability in soil K across the fields. The comparable median and mean K values across most fields indicated minimal outliers. Skewness values were near zero in most fields, with only fields 3, 5, and 7 showing higher skewness (>0.9), suggesting the presence of a few outliers within these datasets. Field 5 had the highest kurtosis (3.92), indicating a more peaked distribution, while other fields had values closer to zero, suggesting more normally distributed data. The wide range observed in soil K highlights significant within-field heterogeneity in K availability. Some of this variability can be attributed to differences in soil texture across fields, as higher clay content and CEC is associated with improved K-holding capacity (Wakeel et al., 2002).

Figure 2.7 demonstrates the soil K variability across the fields, with red or orange indicating areas that require K applications. Classifications of nutrient levels were again based on Auburn University soil test recommendations (Mitchell & Huluka, 2012), and the classifications were as follows: very low (0-34 kg/ha), low (34-67 kg/ha), medium (67-135 kg/ha), high (135-269 kg/ha), and very high (269+ kg/ha). Based on these recommendations, only field 4 had mean soil K concentrations low enough (<120 kg K/ha) to warrant a higher rate of fertilizer application (>50% area applied). Fields 1 and 9 required moderate fertilization requirement, with 25-50% of the field requiring K application. All other fields had significant residual K values, with most requiring application in less than 15% of the field area. Additionally, none of the fields contained areas with very low K levels, and only fields 1 and 9 had areas with low K levels. These values demonstrated that soil K was most frequently above the critical soil test value for yield maximization when compared to soil pH and P.

Table 2.5. Descriptive Statistics for soil K (kg/ha) for the fields used in this study.

Field	Range	Mean	Median	CV (%)	Skewness	Kurtosis
1	39-397	162	165	34	0.39	0.52
2	57-486	206	196	39	0.68	0.32
3	133-390	219	215	19	0.93	1.85
4	64-236	127	125	24	0.62	0.65
5	62-476	179	170	31	1.40	3.92
6	79-272	171	172	22	-0.06	-0.01
7	113-345	190	185	21	1.01	1.93
8	96-344	181	175	25	0.82	0.54
9	57-270	139	140	27	0.43	0.59

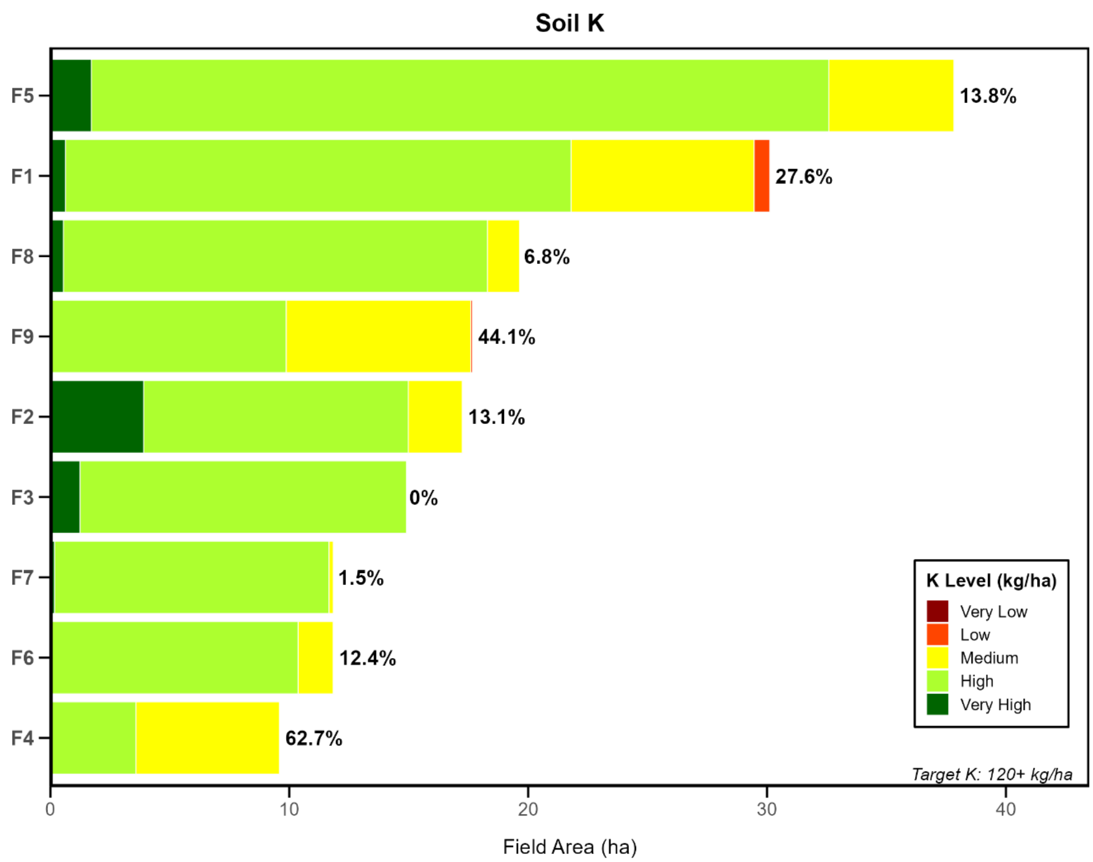

Figure 2.7. Field areas classified in different soil K values. Percentage values represent the total amount of the field that requires fertilizer.

2.4.2 Pearson Correlation Analysis

Pearson correlation analyses suggested no particular trends between the spatial crop and soil properties and nutrient levels within the fields. Therefore, these results are presented and discussed field by field. For this discussion, correlation coefficients (r) ≥0.30 were considered indicative of moderate correlation, whereas values ≥0.60 were considered indicative of strong correlation. All correlation values below 0.30 were considered weak or negligible (Schober et al., 2018). The lack of variability in soil series across fields may be due to the presence of only a few soil types within them. Similarly, correlations with soil series for Field 4 were not possible because only one series was present throughout the field.

Table 2.6. Correlation of various soil and crop spatial properties with soil pH, P and K as per the Pearson Correlation Analysis.

Nutrient	Field	Soil Series	EC Deep	EC Shallow	Elevation	NDVI	Yield
	1	0.12	0.30*	0.19*	0.22*	0.26*	0.30*
	2	-0.17*	–	–	0.37*	0.42*	0.35*
	3	0.14	0.29*	0.19*	-0.05	0.05	-0.02
	4	–	0.46*	–	-0.1	0	–
Soil pH	5	-0.09	0.20*	0.12*	0.05	0.14*	–
	6	0.29*	0.28*	–	0.57*	0.63*	0.69*
	7	0.19*	0.52*	0.39*	0.08	-0.04	-0.04
	8	0.06	0.37*	0.37*	-0.18*	-0.20*	–
	9	-0.1	-0.25*	-0.12	0.03	0.01	–
	1	-0.06	-0.36*	-0.17*	-0.03	-0.25*	-0.03
	2	0.09	–	–	-0.49*	-0.23*	-0.12
	3	-0.17*	-0.30*	-0.48*	-0.28*	-0.44*	-0.25*
	4	–	-0.58*	–	0.25*	-0.15	–
Soil P	5	-0.22*	-0.25*	-0.30*	0.08	-0.06	–
	6	0.35*	-0.01	–	0.03	0.31*	0.26*
	7	0.27*	-0.43*	-0.44*	-0.50*	-0.09	0.14
	8	0.02	0.15*	0.43*	-0.31*	-0.35*	–
	9	-0.30*	-0.30*	-0.36*	0.39*	0.22*	–
	1	-0.03	-0.07	0.15*	0.18*	0.16*	0.29*
	2	0.01	-	-	0.54*	0.50*	0.44*
	3	0.21*	0.33*	0.42*	0.45*	-0.06	0.02
	4	-	0.46*	-	0.13	-0.17	-
Soil K	5	-0.10*	0.49*	0.50*	-0.10*	0.23*	-
	6	0.07	0.41*	-	0.34*	0.50*	0.53*
	7	0.23*	0.32*	0.41*	-0.10	0.08	0.11
	8	-0.11	0.13	0.50*	-0.50*	0.09	-
	9	0.00	0.68*	0.54*	-0.15	0.28*	-

A dash (–) indicates the data type was unavailable for the selected field. * represents a significant relationship at $a = 0.05$.

Correlation coefficients for soil pH were highly variable across fields and spatial data layers, ranging from 0.01 to 0.69. Deep EC showed significant correlations across 7 of 9 fields, with values ranging from weak to strong (r = 0.20 to 0.52). Shallow EC demonstrated significant relationships in six fields with generally weaker coefficients ($r = 0.12$ to 0.39). These findings contrast with Lajili et al. (2021), where the authors evaluated zone delineation based on soil EC data and reported similar r values (0.14-0.15) and no significant correlation between soil pH and EC values. Elevation exhibited significant correlations in five fields, with only two showing moderate correlations ($r = 0.57$, 0.37). Field 6 showed notably strong correlations across multiple predictors, including yield ($r = 0.69$), NDVI ($r = 0.63$), and elevation ($r = 0.57$). Soil series showed significant relationships in only three fields, with $r < 0.30$. Both crop-specific spatial layers (yield and NDVI) showed similar results, with a moderate to strong correlation for the same field. Overall, deep EC and yield retained the strongest predictive potential for soil pH, though performance varied considerably across fields.

Soil P exhibited predominantly negative correlations with predictor variables, with r values ranging from -0.58 to 0.43. Deep EC showed significant correlations in 7 of 9 fields, with most negative and moderate in strength ($r = -0.25$ to -0.58). Field 4 demonstrated the strongest relationship ($r = -0.58$). The negative correlation again contrasted with Lajili et al. (2021) as the authors reported weak (positive) correlations for both shallow and deep EC in their study. Elevation produced significant correlations in six fields, primarily negative and ranging from weak to moderate ($r = -0.30$ to -0.48), with Field 3 showing the strongest relationship ($r = -0.48$). Yield demonstrated significant correlations in only five fields, with mixed results and generally weak coefficients ($r \leq 0.26$). Soil series showed limited predictive ability; significant correlations were found in only five fields, and only two fields had correlations with $r \geq 0.30$. The predominantly negative correlations suggest inverse relationships between soil P levels and most spatial properties, especially soil EC. This could be attributed from phosphorus becoming tied up from the excess salinity that is present in soils with higher EC.

Soil K showed markedly different correlation patterns from soil P, with predominantly positive relationships and r values ranging from -0.50 to 0.68. Deep EC showed significant associations across eight of nine fields, with most being weak to moderate positive correlations (r = 0.32-0.49). Shallow EC showed significant correlations across seven fields with generally strong positive relationships, particularly in Fields 2 and 5 (r = 0.54 and 0.50, respectively). Elevation exhibited significant correlations in seven fields, with Field 8 showing a strong negative correlation (r = -0.50), and a strong correlation with NDVI (r = 0.50). Field 3 showed consistently strong positive correlations across multiple predictors, including elevation (r = 0.42), shallow EC (r = 0.42), and NDVI (r = 0.45). Yield and NDVI both displayed similar correlation values when evaluated on a field-by-field basis. In contrast, soil series showed a low correlation across all fields, with r values mostly ≤ 0.23. The predominantly positive correlations among soil EC, elevation, NDVI, and yield suggest that soil K levels are well associated with these spatial properties and have the most potential for use in the delineation of management zones for precision soil sampling.

2.4.3 Random Forest Modelling

Random forest modelling of each nutrient was performed on a field-by-field basis, using all available spatial data for each field. Similar to the results observed in the Pearson Correlation, model performance varied widely across fields and among soil nutrients. R^2 values ranged from 0.04 to 0.78, indicating a large variation in the model's predictive ability (Table 2.7). Most models for soil pH had lower R^2 values but also very low RMSE, mostly due to the inherently low soil pH values. These results are expected, as Table 2.3 shows CV values of 4.2 to 6.8% for soil pH across all fields. The soil pH model exhibited the lowest performance in Field 7 ($R^2 = 0.04$) and Field 9 ($R^2 = 0.1$), despite having the most spatial data available for these fields. Fields 4 and 6 showed the best predictive ability, with R^2 values of 0.34 and 0.65, respectively, and RMSE values below

0.3. Overall, soil pH was difficult to reliably model due to the inherent low variability in the fields evaluated in this study.

The random forest models for soil P showed improved performance, with R^2 values ranging from 0.16 to 0.78 (Table 2.7). Fields 1, 3, 8, and 9 all had weak values ($R^2 \leq 0.27$), indicating poor model fit in these locations (R^2 = 0.16-0.24). Fields 2 and 4 had better predictive performance, with R^2 values of 0.57 and 0.78, and RMSE values of 22.87 and 8.63, respectively. RMSE was considerably higher for soil P than for soil pH, most likely due to the greater variability of soil P values across the fields.

Soil K models showed the highest consistency across all field models, with R^2 values ranging from 0.28 to 0.76. RMSE values were slightly improved on average relative to soil P, consistent with the CV patterns discussed in the previous section. Fields 2, 8, and 9 had the strongest R^2 values with the model, ranging from 0.50 to 0.76, and RMSE values from 24.12 to 33.47. Field 7 showed the weakest model parameters with an R^2 of 0.28 and RMSE of 28.66. The improved performance of soil K models compared to soil P and pH reiterates the findings from previous sections, showing that soil K variation can be estimated by certain soil and crop spatial properties.

Across all three nutrients of interest, fields 2, 4, and 6 had the most consistent models, producing R^2 values $\geq$ 0.40 in at least two of the three nutrients. The models in Fields 1, 3, and 7 had the lowest overall performance across all three nutrients, with R^2 values below 0.40 for nearly every model.

Table 2.7. Descriptive statistics for random forest models by field.

Field	Model	RMSE	R^2	MAE
1	Soil pH	0.21	0.3	0.17
	Soil P	30.8	0.16	25.37
	Soil K	37.35	0.35	28.94
2	Soil pH	0.43	0.23	0.33
	Soil P	22.87	0.57	18.55
	Soil K	33.47	0.76	28.22
3	Soil pH	0.24	0.16	0.17
	Soil P	31.3	0.27	26.3
	Soil K	28.67	0.38	22.42
4	Soil pH	0.29	0.34	0.22
	Soil P	8.63	0.78	7.25
	Soil K	19.38	0.41	14.65
5	Soil pH	0.25	0.22	0.2
	Soil P	66.79	0.25	51.11
	Soil K	35.71	0.39	27.27
6	Soil pH	0.26	0.65	0.22
	Soil P	14.35	0.33	11.39
	Soil K	27.59	0.46	21.66
7	Soil pH	0.33	0.04	0.27
	Soil P	14.72	0.46	11.87
	Soil K	28.66	0.28	24.01
8	Soil pH	0.28	0.42	0.2
	Soil P	27.41	0.22	18.97
	Soil K	29.99	0.5	24.73
9	Soil pH	0.36	0.1	0.29
	Soil P	18.12	0.24	15.36
	Soil K	24.12	0.65	18.25

Table 2.8 presents the three most strongly correlated spatial layers within each field to different nutrients. Across all fields, elevation, along with deep and shallow EC, were most frequently identified as the strongest contributors to the RF models. The exact values for the variable

importance are presented in Figures A.1, A,2, and A.3 in Appendix A. EC deep was most effective for modelling soil pH data, appearing as a top three predictor in all but one field where it was available. It also showed effectiveness in estimating soil P and K , appearing as a strong predictor in the majority of fields. For all nutrients, Elevation appeared as a significant model contributor across all fields, showing results similar to the Pearson Correlation analysis. This demonstrates the potential of elevation data for zone delineation. Overall, NDVI and yield data had less model contribution, with field 6 showing the most contribution from these datasets. These datasets showed similar results in the previous section, with moderate to high *r* values across all three soil nutrients. Soil series contributed minimally to the models, with it being a top predictor in only two to three fields for each nutrient.

Table 2.8. The top three variables by model importance for each field and nutrient, based on random forest modelling.

Field	Soil pH	Soil P	Soil K
1	Elevation, Yield, EC Deep	EC Deep, NDVI, EC Shallow	Yield, Elevation, NDVI
2	Elevation, Soil Series, NDVI	Soil Series, Elevation, Yield	Soil Series, Elevation, NDVI
3	EC Deep, Elevation, NDVI	EC Shallow, NDVI, Elevation	Elevation, EC Shallow, EC Deep
4	EC Deep, Elevation, NDVI	EC Deep, Elevation, NDVI	EC Deep, NDVI, Elevation
5	EC Shallow, EC Deep, Elevation	EC Shallow, Elevation, Soil Series	EC Shallow, EC Deep, NDVI
6	Yield, Elevation, NDVI	NDVI, Yield, Soil Series	NDVI, Yield, Elevation
7	EC Deep, EC Shallow, Elevation	Elevation, EC Deep, Soil Series	Soil Series, EC Deep, EC Shallow
8	EC Shallow, EC Deep, NDVI	EC Shallow, NDVI, Elevation	EC Shallow, EC Deep, Elevation
9	EC Deep, NDVI, Elevation	Elevation, EC Deep, EC Shallow	EC Deep, NDVI, EC Shallow

2.5 Conclusions

The spatial variability of soil P and K was substantial across all nine fields in this study, with CV values reaching up to 46% for soil P and 39% for soil K, while soil pH data had much less variability, with CV values below 7%. Pearson correlation analysis indicated that no single spatial layer exhibited consistent relationships across all fields, suggesting that a broad approach may not be feasible for zone soil sampling across all fields. Both EC deep and shallow, and elevation were the most correlated layers across nutrients and fields, particularly for soil K, where EC deep showed significant positive associations in eight of nine fields. Soil series showed more limited and inconsistent *r* values, whereas NDVI and Yield were effective in only a single field. Random Forest (RF) modelling confirmed the importance of EC and elevation as the dominant predictors of soil nutrient variability across most fields and nutrients, broadly agreeing with the Pearson correlation results. Model performance varied considerably, with R^2 values ranging from 0.04 to 0.78. Soil K models showed the greatest consistency, while soil pH models were consistently the weakest, likely due to the inherently low spatial variability of pH data across all fields in this study. The results from the RF models supported the use of soil EC (deep and shallow) and elevation as the primary spatial datasets for delineating management zones, with NDVI and yield being possible predictors in certain fields. Based on both correlation analysis and RF modelling, a combination of soil EC and elevation is recommended as the most practical and broadly applicable basis for management zone delineation within the Southeast. However, the lack of a universally predictive layer set across all nine fields suggests that zone delineation strategies should be evaluated on a field-by-field basis rather than a similar broad approach across the whole farm or region.

Chapter Three

Delineation and Validation of Management Zones Based on Soil and Crop Spatial Properties for Precision Soil Sampling

3.1 Abstract

Precision soil sampling plays a critical role in site-specific nutrient management in precision agriculture. Grid sampling has become well established across the Southeast for this purpose. Zone-based soil sampling has gained interest as a cost-effective alternative to grid sampling for site-specific nutrient management in the southeastern United States, yet limited research exists on effective management zone (MZ) delineation strategies for this region. This study evaluated four spatial data layer combinations for MZ delineation across four fields (11.8–30.1 ha) in the Coastal Plain region of Georgia. Delineation strategies included EC Only (S1), EC + Elevation (S2), EC + Elevation + NDVI (S3), and EC + Elevation + Yield (S4), with zones generated using k-means clustering at k = 3, 4, and 5. Zone performance was assessed using variance reduction (VR) of soil pH, phosphorus (P), and potassium (K) relative to a high-density 0.1-ha grid soil sampling baseline, and zone means were subjected to an ANOVA and Tukey's post hoc test. Results indicated that four zones consistently provided the best balance between VR improvement and practical implementation, with five zones offering only marginal gains over four. S1 performed the worst across all fields and zone counts, while S2 showed competitive performance with the three-layer strategies across most fields. S3 and S4 produced notably higher VR in one field containing an area of recently cleared land, where vegetative indices responded to the strong spatial differences. Soil K showed the greatest zone differentiation across all fields, while soil pH was the most difficult nutrient to differentiate due to its inherently low spatial variability. Overall, Strategy

2 was identified as the most broadly applicable and practically efficient delineation strategy, offering improved performance over EC alone while minimizing data requirements.

3.2 Introduction

Agricultural fields in the southeastern United States are characterized by considerable spatial nutrient variability, driven by factors such as soil type, topography, and previous management practices(Beckett & Webster, 1971). This inherent variability often contributes to inefficiencies in fertilizer and lime applications, leading to under- and over-application of nutrients and soil amendments across most of the fields. To combat these issues, variable-rate (VR) applications of fertilizer and agricultural lime are implemented, enabling growers to align nutrient requirements with localized soil conditions and effectively manage in-field nutrient variability. Proper implementation of VR applications relies on precision soil sampling to obtain information about existing nutrient variability within the agricultural fields. Grid and zone sampling are two main types of precision soil sampling strategies utilized across the United States, with grid sampling being predominantly used across approximately two-thirds of the fields sampled each year acreage (McFadden et al., 2023). This wide adoption stems from a variety of benefits, with ease of implementation and accurate spatial nutrient representation being the primary benefits (Ackerson, 2018). However, these benefits are often realized with high-resolution grids between 0.4 and 1.01 ha (1-2.5 acres) in size, which can create a significant increase in labor and lab analysis (Mallarino & Wittry, 2004). An alternative to grid soil sampling, zone-based soil sampling, has the ability to retain the spatial representation of high-density grids, but reduces both labor and analytical costs from a reduction in sample number (Valente et al., 2024). While the implications of various grid sizes and sampling methods have been explored, there remains a considerable knowledge gap regarding the effective implementation of zone soil sampling, particularly in using various spatial data layers and delineation methodologies.

Management zones (MZs) are sub-field areas that share similar soil or crop characteristics and are frequently used to inform site-specific management of crop inputs; however, their adoption remains low. As VR technology becomes more widely adopted (McFadden et al., 2023), the interest in using MZs for precision soil sampling has increased, especially in the southeastern US, where soil variability is often pronounced and challenging to capture with uniform soil sampling approaches. MZ-based sampling strategies have the potential to reduce sampling costs compared with intensive grid sampling (0.4-1.01 ha grids), provided that the zones are homogeneous and accurately represent nutrient variability within the fields. The use of MZs in agriculture dates back more than 50 years, with early delineation methods relying heavily on historical knowledge and visual observations (Fleming et al., 2000). As technologies and methods for rapid access to spatial data have advanced, interest in using soil and crop properties for management zones has also increased.

Several studies have evaluated MZs based on various soil or crop spatial attributes, such as historical yield (Hornung et al., 2006; Jaynes et al., 2005), soil type (Coleman, 2021), elevation (Schepers et al., 2004), organic carbon (Gozdowski et al., 2014), and soil electrical conductivity (Lajili et al., 2021). Although management zone approaches based on a specific soil or crop spatial property can be effective, utilizing multiple relevant spatial data layers can often produce zones that better reflect within-field nutrient variability (Pusch et al., 2023). However, studies investigating such MZ delineation approaches remain limited, with recent research in the southeastern US focused primarily on using only soil EC or grower-generated (often hand-drawn) zones for soil sampling. Additionally, the majority of available research focuses on using MZs for site-specific irrigation and nitrogen management rather than for managing nutrients such as soil P and K (Haghverdi et al., 2015; Taylor et al., 2007; Thompson et al., 2004; Tucker, 2024).

Several challenges exist when evaluating spatial data layers for MZ delineation. One of the most critical issues is the collection of reliable spatial datasets for use in delineation methods. The quality of the data can vary between sources, and it may often lead to less effective management zones (Gozdowski et al., 2014). While yield and topographic data can be collected directly from modern machinery, their accuracy is often constrained by sensor calibration, machine delays, and GPS precision. Soil EC and elevation data tend to be more reliable but still require appropriate post-processing to remove outliers and ensure that spatial patterns represent true differences in soil properties. The effectiveness of zone sampling also depends heavily on homogeneity within the zone; thus, selecting data layers that correlate strongly with soil nutrient availability is essential (Javadi et al., 2022).

In Summary, precision soil sampling remains a critical component of site-specific nutrient management, and soil sampling using 0.4- to 1.0-hectare grids is commonly recommended to capture spatial variability. However, grid sampling can be expensive and labor-intensive across large acreages, creating a financial barrier for many growers. Zone-based soil sampling offers a more economical alternative if MZs can accurately reflect in-field nutrient levels. Limited information exists on evaluating crop and soil properties for delineating management zones for soil sampling and their effectiveness in depicting nutrient variability within agricultural fields. Therefore, a comprehensive study across the southeastern US, using high-resolution spatial data and validated nutrient measurements, was undertaken to identify effective approaches for delineating MZs and provide guidance on the suitability of different soil and crop attributes for zone-based soil sampling. The main objective of this study was to delineate management zones for soil sampling using various spatial data layers and to evaluate their ability to capture the true variance in soil pH, phosphorus (P), and potassium (K) within agricultural fields in the southeastern US.

3.3 Materials and methods

3.3.1 Experimental Sites

The data for this study were gathered in 2024 and 2025 across 4 fields in the Southeastern US. Selected fields ranged in size from 11.8 to 30.1 ha (Table 3.1). All fields were located in the Coastal Plain region of the Southeast and contained similar soil types. Soil series information for these fields is shown in Figure 3.1. Fields were selected based on grower knowledge of higher variability in soil pH, P, and K in fields across the region. Field rotations consisted of cotton-cotton-peanut, with reduced (strip) tillage in cotton and conventional (bedding) tillage for peanut.

Table 3.1. Information on the fields used in this study.

Field	Longitude	Latitude	Size (ha)	Water Management
1	-83.4944	32.16946	30.1	Irrigated
2	-83.902	32.06156	14.9	Irrigated
3	-83.9015	31.69454	11.8	Irrigated
4	-83.4668	32.18166	11.9	Dryland

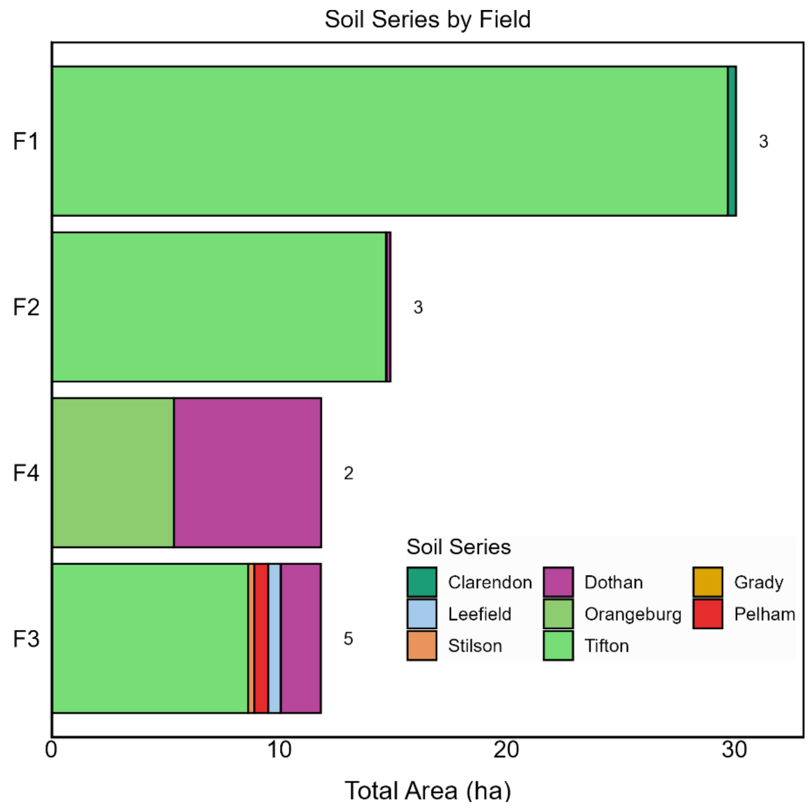

Figure 3.1. Soil series information by area for each field. Numbers after bars represent # of distinct soil series within a field

3.3.2 Soil Sampling and Data Acquisition

Different types of soil and crop spatial data, including soil EC, elevation, NDVI, and yield, were collected for the selected fields. Soil electroconductivity (EC) data were collected across all four fields using a Veris Technologies Mobile Sensor Platform (MSP) 3100 (Salina, KS) pulled with a UTV. The UTV was equipped with a Trimble GPS 7500 auto-guidance system (Trimble, Westminster, CO) with Real-Time Kinematics (RTK) capabilities, enabling georeferencing of spatial data with an accuracy of ±1 cm. Both shallow (30 cm) and deep (90 cm) soil EC data were collected in each field, except in Field 3, where only deep EC was collected due to equipment limitations. The sensor was set to collect one data point every second (1 Hz), and data was collected by making consecutive passes spaced 12 m apart within each field. The raw EC data was processed using Field Fusion software from Veris Technologies (Salina, KS) to remove any erroneous data and other outliers. In addition to shallow and deep EC data, the sensor also recorded elevation across the field at the same spatial resolution. Yield data was available for all four fields in this study and collected using a John Deere CP 770 cotton picker equipped with an onboard yield monitor. Raw data was uploaded directly into the Accufield software (GreenPoint Ag, Decatur, AL). Only fields that had two or more years of available yield data were used to create a stable normalized yield dataset. When cleaning the yield data, the procedure outlined in Sudduth et al. (2012) was followed to ensure that the data were normalized for the analysis.

Normalized difference vegetation index (NDVI) data was accessed using the AgLeader SMS Advanced software (Ag Leader Technology, Ames, IA), which utilizes Sentinel-2 satellite imagery at a spatial resolution of 10 × 10 m to generate this crop index. The NDVI data was downloaded in August 2024, as most fields were at or near peak growth and canopy coverage during this time. The maximum allowed cloud cover was set at 10%; however, the maximum cloud cover downloaded for any field was 2% or less. Any non-vegetative or cirrus cloud cover pixels were removed, but no more than 10% of the total pixels were removed in any field. Table 3.2 provides information on the various types of data collected and utilized for each field.

Table 3.2. Various spatial soil and crop properties used for management zone analysis.

Field	1	2	3	4
Variable (Attribute)				
EC Shallow	X	X	-	X
EC Deep	X	X	X	X
Elevation	X	X	X	X
Normalized Yield	X	X	X	X
NDVI	X	X	X	X
Soil Type	X	X	X	X

3.3.3 Grid Soil Sampling

A high-density grid soil sampling was carried out across all 4 fields to capture the nutrient variability within each field. For this soil sampling, field boundaries were loaded into AccuField software (GreenPoint Ag, Decatur, AL), and a 32 × 32 m (0.1 ha) sampling grid was generated across each site. These grids were transferred to the FieldAlytics mobile application running on a tablet and paired with a BadElf GPS Pro receiver, which provides 2.5 cm accuracy for in-field navigation to each grid. At every grid point, 12 to 15 soil cores were extracted to a depth of 15 cm using a hand probe and combined to produce a single composite sample per grid cell. Labeled samples were shipped to the University of Georgia Agricultural and Environmental Services Laboratories (AESL, Athens, GA) for analysis. Extractable phosphorus (P) and potassium (K) were determined via Mehlich-1 extraction, while soil pH was measured in a 0.01 M $CaCl_2$ suspension (Kissel & Sonon, 2011). The resulting high-resolution nutrient maps served as a ground-truth reference based on which the performance of MZs was evaluated.

3.3.4 Management Zone (MZ) Delineation

MZs were generated through k-means clustering using four combinations of the available spatial layers: (Strategy 1) Soil EC Only; (Strategy 2) Soil EC + Elevation, which combined the EC layers with gathered elevation data; (Strategy 3) Soil EC + Elevation + NDVI; and (Strategy 4) Soil EC + Elevation + Yield, which substituted multi-year normalized yield for NDVI. These delineation strategies will be referred to as S1, S2, S3, and S4 from here forward. All spatial layers were centered and scaled prior to clustering to prevent any single variable from causing skewness due to scaling. Clustering was executed in R (RStudio 2025.05.1) with the *kmeans* function. Each layer or combination was clustered at k = 3, 4, and 5, yielding 12 zone maps per field and 48 configurations in total. The number of zones chosen was based on the average silhouette width computed with the *cluster* package (Maechler et al., 2022), and MZ counts were limited to values that were practically implementable for available software and field equipment. The k value yielding the highest mean silhouette width was recorded as the analytically optimal cluster count, although zones were ultimately evaluated at k = 3, 4, and 5 because most combinations fell within this range and remained practical.

3.3.5 Management Zone (MZ) Validation

To examine the effectiveness of the MZ delineation methods, two evaluations were conducted: total variance reduction was calculated, and the higher-performing strategies were compared, resulting in a one-way Analysis of Variance to determine whether zone differences were meaningful. Variance reduction was calculated similarly to Miao et al. (2018), as follows:

$$VR\ (\%) = [(Var_f - Var_z / Var_f] \times 100$$

Where, Var_f = variance of nutrients across the full field

Var_z = variance of nutrients within the zone

VR = variance reduction of nutrients expressed in percentage

Positive VR values indicate that zones are internally more uniform than the field. Negative VR values signal that the clustering-imposed boundaries increased within-zone heterogeneity relative to the undivided field. VR was computed individually for soil pH, P, and K, and the mean of the three nutrients was also reported as a composite index of overall zone quality. A one-way ANOVA was performed on these strategies within the zone configurations to test whether zone membership explained a significant portion of the variance in nutrient levels. Where the overall F-test was significant, Tukey's post hoc test was performed with $\alpha = 0.05$ to identify which zone pairs differed. All mean-separation analyses were performed with the *agricolae* package in R (de Mendiburu, 2021).

3.4 Results and Discussion

3.4.1 Optimal Number of Zones

Field 1

Field 1 was the largest site in the study (30.1 ha) and had a range of optimal clustering values. S1 produced silhouette widths that remained relatively stable across all k values, ranging from approximately 0.50 to 0.56, with the optimal k occurring at k = 10 (Figure 3.3). The relatively flat silhouette curve for EC only in this field indicates that the EC data maintained well-separated clusters regardless of the number of zones created. S2 peaked at k = 2, while both three-layer strategies peaked at k = 4, with EC + Elevation + NDVI reaching approximately 0.37 and EC + Elevation + Yield reaching approximately 0.35. Only S3 and S4 showed a significant reduction in silhouette widths.

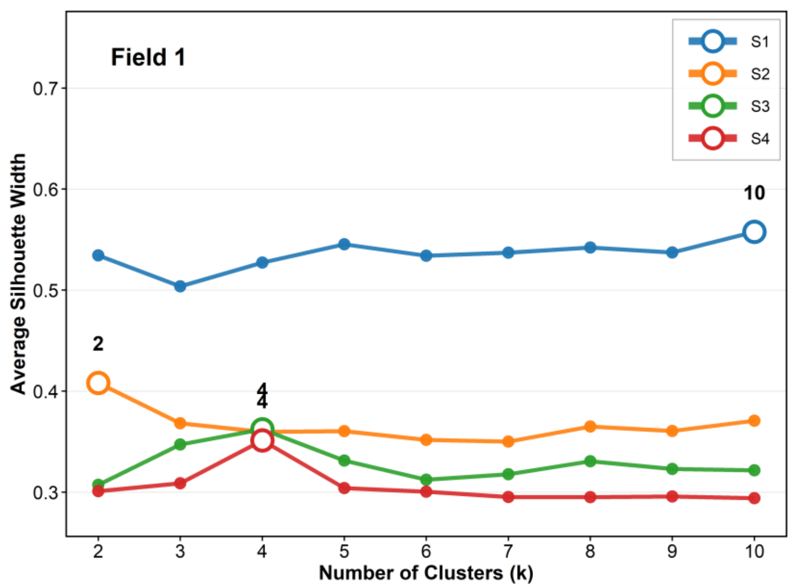

Figure 3.3. Optimal cluster counts for each delineation strategy for Field 1.

Field 2

Field 2 (14.9 ha) also had significant separation between S1 and all other strategies (Figure 3.4). S1 reached its maximum silhouette width of 0.61 at k = 6, with a pronounced increase from k = 2 through k = 6 before declining slightly at higher k values. S2 peaked at k = 2, while S3 peaked at k = 5 and S4 at k = 2. The yield-based strategy (S4) had the lowest silhouette widths overall, declining below 0.30 at higher k values, indicating that the addition of yield introduced spatial complexity that weakened the datasets' ability to be clustered separately into two zones.

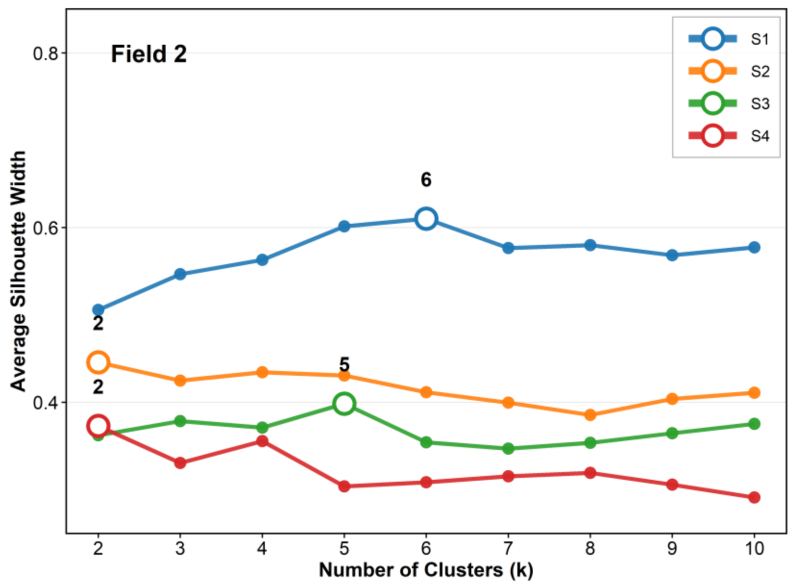

Figure 3.4. Optimal number of clusters for each delineation strategy for Field 2.

Field 3

Field 3 (11.8 ha) produced the highest EC Only silhouette width of any field, peaking at k = 2 with a value of approximately 0.71 (Figure 3.7). Increasing the number of clusters showed a consistent diminishing return in cluster separation. The strong two-cluster structure in EC data for this field suggests two well-separated areas, likely because one area of the field contains recently cleaned-up farmland. The multi-layer strategies all peaked at k = 3, with S2 reaching approximately 0.45, S3 approximately 0.40, and S4 approximately 0.38.

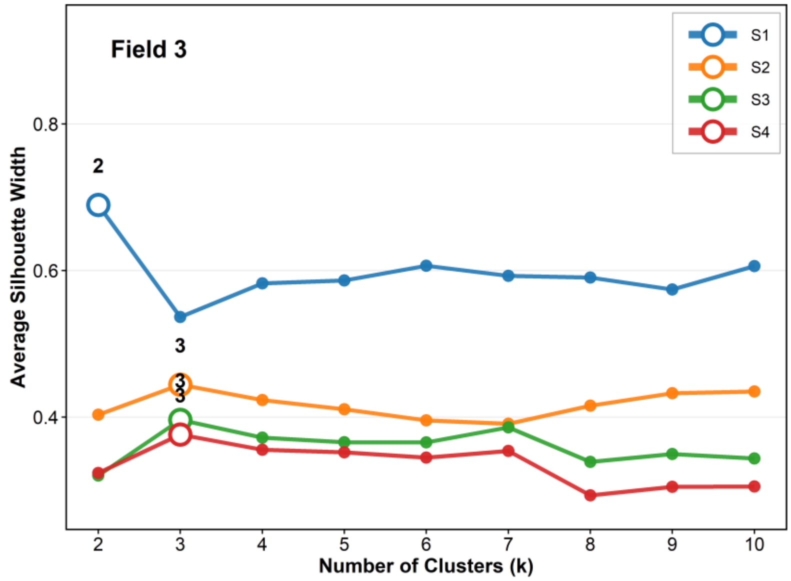

Figure 3.5. Optimal number of clusters for different MZ delineation strategies for Field 3.

Field 4

Field 4 (11.9 ha) was the only dryland site and showed a distinctive pattern: all four strategies achieved their highest silhouette widths at k = 2 (Figure 3.6). S1 peaked at approximately 0.70, EC + Elevation at approximately 0.51, EC + Elevation + NDVI at approximately 0.49, and EC + Elevation + Yield at approximately 0.45. Silhouette widths declined steadily as k increased for all strategies. This presents a scenario similar to field 3, with two distinct areas in the field and less benefit from more than 2 zones. S1 had a much more similar pattern to S2, S3, and S4 than to some of the previous datasets. With nearly every optimal zone number falling between 2 and 6, it was decided to create zones in counts of 3, 4, and 5 to evaluate most of the optimal clustering counts, while also keeping clusters large enough to be treated as individual areas in the field.

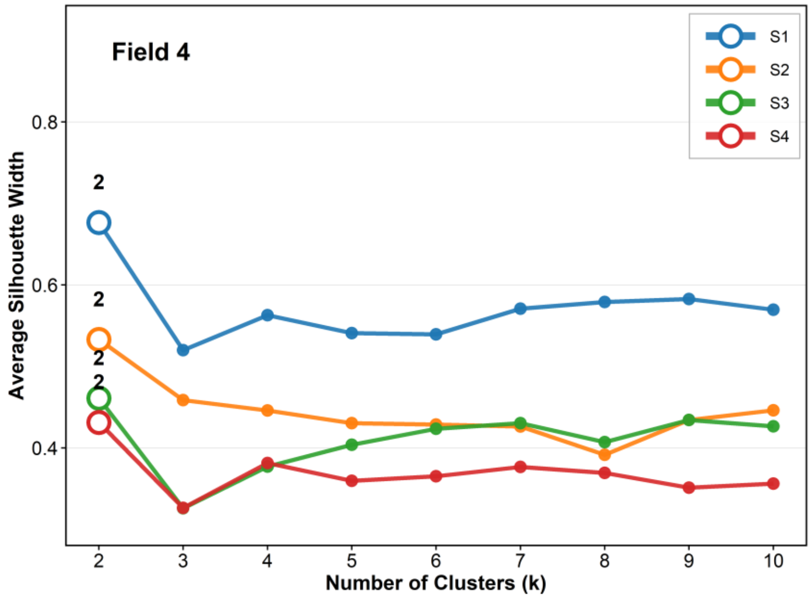

Figure 3.6. Optimal number of clusters for different MZ delineation strategies for field 4.

3.4.2 Variance Reduction

The number of management zones (3, 4, or 5) affected performance across all strategies. For strategy 4, 5 clusters produced the highest mean variance reduction (VR) across all nutrients (20.33%), followed by 4 clusters (19.20%), and 3 clusters (13.72%) (Table 3.3). This trend was consistent across all four zone delineation methods, with mean VR values increasing as the cluster count rose from 3 to 5. However, diminishing returns were observed as cluster count increased. The improvement in mean VR from 3 to 4 clusters averaged 3.7% across all strategies and nutrients, while the increase from 4 to 5 clusters averaged only 2.3%. For strategy 1, only a 1.14% increase was observed from 3 to 4 clusters and a 2.36% from 4 to 5 clusters, suggesting that this strategy may not benefit substantially from increased zone complexity.

Table 3.3. Mean variance reduction (%) averaged across nutrients by field.

Field	k	S1	S2	S3	S4
1	3	7.78	6.29	2.27	4.52
	4	9.90	7.94	7.98	9.22
	5	10.39	11.38	10.10	11.30
2	3	8.93	9.38	9.29	13.82
	4	12.42	15.62	16.45	13.60
	5	13.11	14.91	11.78	14.26
3	3	9.55	13.09	19.73	23.64
	4	11.22	11.81	23.78	32.54
	5	17.11	14.46	37.22	32.21
4	3	17.38	15.98	15.72	12.88
	4	14.67	20.68	15.48	21.44
	5	17.01	25.27	18.29	23.56

In terms of site-specific analysis, across all four fields, VR values increased with additional zones but with diminishing returns. The improvement from k = 3 to k = 4 averaged 3.7 percentage points across strategies, while the gain from k = 4 to k = 5 averaged only 2.4 percentage points. S2, S3, and S4 generally outperformed S1, with the advantage being greatest in field 3, where a large difference between the two datasets was seen. However, in fields where these correlations were weaker (Field 1), S1 still showed a similar VR reduction at lower MZ counts. One takeaway from this data is that yield or NDVI data is particularly useful when evaluating fields containing newly cropped ground.

Table 3.4 Mean values for soil pH, P and K within each zone in field 1 (k = 5).

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	149 b	79 c	6.6 a
	2	142 b	127 a	6.1 c
	3	174 a	98 ab	6.6 a
	4	168 ab	117 a	6.5 b
	5	145 b	85 bc	6.6 ab
S2	1	122 c	69 d	6.6 a
	2	173 a	97 bc	6.6 a
	3	167 ab	87 cd	6.6 a
	4	171 ab	114 ab	6.6 a
	5	140 bc	119 a	6.3 b
S3	1	166 ab	89 b	6.6 a
	2	180 a	111 a	6.6 a
	3	152 bc	76 b	6.6 a
	4	127 c	117 a	6.3 b
	5	154 abc	121 a	6.5 a
S4	1	128 b	79 b	6.6 a
	2	174 a	100 a	6.6 a
	3	170 a	82 b	6.6 a
	4	133 b	124 a	6.2 b
	5	170 a	116 a	6.6 a

Means with the same letters represent zones that were not significantly different from each other for that nutrient in the respective field. $\alpha = 0.05$

Field 2 at k = 4 presented the most challenging case for zone differentiation (Table 3.5). While K-means showed clear separation under S3, the mean values for soil pH were not statistically significant across all four zones for most strategies. This confirms the VR findings showing that the narrow soil pH range in Field 2 prevented meaningful zone-based differentiation for this nutrient.

Table 3.5. Mean values for each zone in field 2 (k = 4).

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	279 a	118 b	6.3 ab
	2	176 c	169 a	6.2 b
	3	220 b	113 b	6.5 a
	4	224 b	128 b	6.4 ab
S2	1	242 a	114 b	6.4 ab
	2	237 a	123 b	6.4 b
	3	187 b	164 a	6.3 b
	4	200 b	118 b	6.5 a
S3	1	288 a	118 b	6.2 a
	2	194 c	121 b	6.4 a
	3	200 c	186 a	6.4 a
	4	233 b	119 b	6.4 a
S4	1	295 a	120 b	6.3 a
	2	187 c	183 a	6.3 a
	3	199 c	123 b	6.5 a
	4	234 b	121 b	6.4 a

Means with the same letters represent zones that were not significantly different from each other for that nutrient in the respective field. $\alpha = 0.05$

Field 3 at k = 4 using strategy 4 produced the clearest zone differentiation of any configuration (Table 3.6). All four zones differed significantly in soil P, K, and pH simultaneously. Mean values for K across the zones ranged from 126 to 199 kg/ha for K, 21 to 48 kg/ha for P, and 5.6 to 6.6 for soil pH. In contrast, S1 produced zones that were not significantly different for P, further demonstrating the usefulness of an additional vegetative dataset for MZ delineation.

Table 3.6 Mean values for each zone (k = 4) in field 3.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	154 b	38 a	6.1 b
	2	231 a	30 a	6.3 ab
	3	173 b	44 a	6.5 a
	4	181 ab	39 a	6.5 a
S2	1	178 ab	44 a	6.6 a
	2	202 a	35 a	6.6 ab
	3	154 c	38 a	6.1 c
	4	164 bc	44 a	6.2 bc
S3	1	179 ab	43 a	6.6 a
	2	200 a	34 ab	6.6 a
	3	100 c	16 b	5.5 c
	4	162 b	44 a	6.2 b
S4	1	181 ab	44 ab	6.6 a
	2	126 c	21 c	5.6 c
	3	165 b	48 a	6.3 b
	4	199 a	34 bc	6.5 ab

Means with the same letters represent zones that were not significantly different from each other for that nutrient in the respective field. $\alpha = 0.05$

Field 4 at k = 5 produced moderate zone differentiation across all four delineation strategies, with the most significantly different zones for soil K and the least difference in soil pH (Table 3.7). Minimal differences were observed for soil P and pH overall. S4 produced the strongest overall differentiation in soil K, with Zones 1, 3, and 5 significantly higher than Zones 2 and 4, and also yielded three distinct pH groupings across the five zones. Despite these differences among strategies, soil pH separation remained limited across all four configurations, consistent with the narrow pH range observed in the other fields evaluated in this study.

Table 3.7 Mean values for each zone (k = 5) in field 4.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	197 ab	45 ab	6.4 ab
	2	180 b	55 a	6.3 b
	3	243 a	25 b	6.7 a
	4	187 ab	62 a	6.2 b
	5	205 ab	33 b	6.7 a
S2	1	208 a	23 c	6.6 a
	2	195 ab	42 b	6.3 b
	3	214 a	49 b	6.6 a
	4	166 b	47 b	6.2 b
	5	188 ab	67 a	6.3 b
S3	1	184 ab	72 a	6.6 ab
	2	174 ab	53 a	6.2 c
	3	205 a	57 a	6.4 ab
	4	170 b	45 a	6.3 bc
	5	206 a	22 b	6.6 a
S4	1	216 a	25 c	6.6 a
	2	169 b	57 ab	6.3 c
	3	204 a	45 b	6.5 ab
	4	168 b	46 b	6.2 c
	5	206 a	65 a	6.4 bc

Means with the same letters represent zones that were not significantly different from each other for that nutrient in the respective field. $\alpha = 0.05$

Across all four fields, the ability to produce statistically different zones depends on both the inherent nutrient variability and the spatial data layers used for delineation. Fields with higher nutrient variability (Field 3 for soil pH and Field 4 for soil P) consistently produced stronger zone separation, while low-variability nutrients (soil pH in Field 2) had little differentiation in zone values. Second, multi-layer strategies generally improved zone differentiation over S1, particularly for nutrients that did not correlate strongly with soil EC alone, such as soil pH.

3.5 Conclusions

Interest in zone-based soil sampling strategies has rapidly increased with the growing availability of spatial datasets in agriculture across the Southeastern United States. The current standard for soil sampling among most growers is grid sampling, which is effective for evaluating nutrient variability but also increases labor and sample analysis costs with higher-density grids. This study was performed to evaluate the zone-based sampling alternative to this sampling method. Given the limited research available on zone sampling in the region, this study was conducted to evaluate the potential of specific data layers to generate management zones for precision soil sampling. Four fields were evaluated based on their representation of the region, with some irrigation and newly cropped ground present. When considering overall nutrient variance reduction, a higher number of management zones (4 or 5) was more effective than utilizing only 3 zones. Additionally, strategy 1 (EC Only) had the worst performance across all fields and MZ counts. When combined with elevation data (S2), performance increased across the majority of fields, and was often comparable to S3 and S4. Exceptional performance was observed for S3 and S4 in field 3, which contained an area of new ground that directly affected the vegetative indices used in this dataset, resulting in much higher VR percentages. When evaluating mean differences for the highest-performing datasets and MZ count combinations, soil pH values often showed the greatest similarity due to the inherent low variability in these datasets. Soil K often had the most distinct zones across all fields, given that it is the most variable nutrient within the dataset. Outside of field three, S2, S3, and S4 all showed similar performance in terms of mean separation of all three nutrients. Based on this analysis, S2 had the most promising performance across all fields while also keeping MZ delineation simpler. Based on the variance-reduction numbers, 4 management zones were the most effective at maintaining higher VR percentages, with 5 zones showing only marginal improvement over 4, and 3 zones being the least effective across all fields.

Chapter Four

Conclusions

Implementing precision zone soil sampling techniques is often more complex than grid soil sampling. With limited research available on the subject, studies were conducted to evaluate its potential in the Southeast, with the first focused on identifying spatial data layers that best represent soil pH, P, and K variability, and the second on evaluating the performance of various MZ delineation strategies and zone counts. Overall, the results of these studies provide practical guidance for growers and agronomic consultants seeking to implement zone-based soil sampling as an alternative to high-density grid sampling.

In the first study, soil P and K exhibited substantial spatial variability across the nine fields evaluated, with CV values reaching as high as 46% for P and 39% for K, while soil pH showed considerably less variability, with CV values consistently below 10%. Pearson correlation analysis indicated that no single spatial data layer was consistently predictive of nutrient variability across all fields, suggesting that a broad, uniform approach to MZ delineation was less feasible. EC deep and EC shallow were the most broadly correlated layers, particularly for soil K, where deep EC showed significant positive associations in eight of the nine fields. Soil series showed the weakest relationship with all three nutrients, while NDVI and yield data were strongly correlated only in a few cases. Random Forest (RF) modeling confirmed that EC and elevation are the dominant predictors of soil nutrient variability across most fields and nutrients. Model performance varied considerably, with R^2 values ranging from 0.04 to 0.78. Soil K models showed the greatest consistency across fields, while soil pH models were consistently the weakest, likely due to the inherently low spatial variability of this nutrient. Based on both correlation analysis and RF modeling, a combination of soil EC and elevation is recommended as the most practical and broadly applicable basis for MZ delineation, with NDVI and yield serving as potentially useful additions on a field-by-field basis.

The second study evaluated four delineation strategies – EC Only (S1), EC + Elevation (S2), EC + Elevation + NDVI (S3), and EC + Elevation + Yield (S4) – across four fields using k-means clustering at zone counts of three, four, and five, with performance assessed against a high-density 0.1-ha grid soil sampling baseline. S1 had the worst performance across all fields and zone counts, confirming that EC alone captures an insufficient portion of within-field nutrient variability for effective zone-based sampling. When elevation was added to EC in S2, performance increased across most fields and was often comparable to that of the three-layer strategies, S3 and S4. The most notable exception was Field 3, which contained an area of new ground that introduced pronounced yield differences, making vegetative indices more effective in this location. Substantially higher VR values were observed for S3 and S4, with S4 reaching up to 32.54% at $k = 4$ and S3 reaching 37.22% at $k = 5$ in that field. Outside of Field 3, the addition of NDVI or yield data provided limited improvement over S2 in terms of both variance reduction and mean zone separation. These findings suggest that S2 represents the most broadly applicable and practically efficient delineation strategy, offering improved performance over S1 while reducing data requirements relative to S3 and S4.

These studies were conducted across a limited number of fields within the Coastal Plain region of Georgia, and expanding this research to more fields across a broader range of soil types, field sizes, topographic conditions, and cropping systems would help develop a more general approach to MZ delineation. Future research should also investigate additional spatial data layers, with many more layers achievable with elevation data alone. While this study used NDVI imagery acquired during peak growth, evaluating more temporally variable datasets could also yield greater predictive power. Additionally, alternative clustering approaches beyond k-means, such as fuzzy c-means clustering or other unsupervised classification methods, could be explored to determine whether these techniques yield zones more effective than those found in this study. Evaluating the temporal stability of these zones over multiple years would also provide valuable insight into their effectiveness over time.

References

Ackerson, J. (2018). Soil Sampling Guidelines.

Agleader SMS. (2004). https://www.agleader.com/sms-software/ Ag Leader Technology, Ames, IA.

Amer, B. S., Moussa, K. F., Sheha, A. A., & Abdel-Fattah, M. K. (2021). Delineation of Site-Specific Management Zones Using Multivariate Analysis and Geographic Information System Technique. *Plant Archives*, 21(1), Article No 1. https://doi.org/10.51470/PLANTARCHIVES.2021.v21.no1.185

Asare, E., & Segarra, E. (2018). Adoption and extent of adoption of georeferenced grid soil sampling technology by cotton producers in the southern US. *Precision Agriculture*, 19(6), Article 6. https://doi.org/10.1007/s11119-018-9568-3

Baxter, S. J., & Oliver, M. A. (2005). The spatial prediction of soil mineral N and potentially available N using elevation. *Geoderma*, 128(3–4), 325–339. https://doi.org/10.1016/j.geoderma.2005.04.013

Beckett, P. H. T., & Webster, R. (1971). Soil Variability: A Review. CAB International, 34(1).

Boydell, B., & McBratney, A. B. (2006). Identifying Potential Within-Field Management Zones from Cotton-Yield Estimates. *Precision Agriculture*, 3, 9–23.

Buscaglia, H. J., & Varco, J. J. (2003). Comparison of Sampling Designs in the Detection of Spatial Variability of Mississippi Delta Soils. Soil Science Society of America Journal, 67(4), 1180–1185. https://doi.org/10.2136/sssaj2003.1180

Castrignanò, A., Buttafuoco, G., Comolli, R., & Castrignanò, A. (2011). Using Digital Elevation Model to Improve Soil pH Prediction in an Alpine Doline. Pedosphere, 21(2), 259–270. https://doi.org/10.1016/S1002-0160(11)60126-4

Climate Corporation. (2021). Maximize Results with Our Digital Farming Solution | Climate FieldView. https://climate.com/en-us.html

Coleman, A. M. (2021). Development of a Management Zone Scoring Index & Economic Returns of Grid & Management Zone Soil Sampling in Coastal Plain Soils of the Southeastern United States. Clemson University.

Dangal, S. R. S., Sanderman, J., Wills, S., & Ramirez-Lopez, L. (2019). Accurate and Precise Prediction of Soil Properties from a Large Mid-Infrared Spectral Library. Soil Systems, 3(1), 11. https://doi.org/10.3390/soilsystems3010011

Dass, A., Suri, V., & Choudhary, A. (2014). Site-Specific Nutrient Management Approaches for Enhanced Nutrient-Use Efficiency in Agricultural Crops. Research & Reviews: Journal of Crop Science and Technology, 3, 1–6.

de Mendiburu, F. (2021). agricolae: Statistical Procedures for Agricultural Research. https://CRAN.R-project.org/package=agricolae

Ditzler, C. (2017). A Glossary of Terms Used in Soil Survey and Soil Classification.

Duffera, M., White, J. G., & Weisz, R. (2007). Spatial variability of Southeastern U.S. Coastal Plain soil physical properties: Implications for site-specific management. Geoderma, 137(3–4), Article 3–4. https://doi.org/10.1016/j.geoderma.2006.08.018

Escadafal, R. (1993). Remote sensing of soil color: Principles and applications. Remote Sensing Reviews, 7(3–4), 261–279. https://doi.org/10.1080/02757259309532181

Farmaha, B., Caughman, W., & Park, D. (2020). Precision Agriculture-Based Soil Sampling Strategies. Clemson Cooperative Extension, Land-Grant Press by Clemson Extension.

Fixen, P. E. (2020). A brief account of the genesis of 4R nutrient stewardship. Agronomy Journal, 112(5), 4511–4518. https://doi.org/10.1002/agj2.20315

Fleming, K. L., Westfall, D. G., Wiens, D. W., & Brodahl, M. C. (2000). Evaluating Farmer Defined Management Zone Maps for Variable Rate Fertilizer Application. Precision Agriculture, 2(2), 201–215. https://doi.org/10.1023/A:1011481832064

Flowers, M., Weisz, R., & White, J. G. (2005). Yield-Based Management Zones and Grid Sampling Strategies: Describing Soil Test and Nutrient Variability. Agronomy Journal, 97(3), Article 3. https://doi.org/10.2134/agronj2004.0224

Gallardo-Romero, D. J., Apolo-Apolo, O. E., Martínez-Guanter, J., & Pérez-Ruiz, M. (2023). Multilayer Data and Artificial Intelligence for the Delineation of Homogeneous Management Zones in Maize Cultivation. Remote Sensing, 15(12), Article 12. https://doi.org/10.3390/rs15123131

Gozdowski, D., Stępień, M., Samborski, S., Dobers, E. S., Szatyłowicz, J., & Chormański, J. (2014). Determination of the Most Relevant Soil Properties for the Delineation of Management Zones in Production Fields. Communications in Soil Science and Plant Analysis, 45(17), 2289–2304. https://doi.org/10.1080/00103624.2014.912289

Greenpoint Ag. (2021). FieldAlytics. https://accufield.agri-afc.com/cgi-bin/login.py

Guo, W., Maas, S. J., & Bronson, K. F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. Precision Agriculture, 13(6), Article 6. https://doi.org/10.1007/s11119-012-9277-2

Haghverdi, A., Leib, B. G., Washington-Allen, R. A., Ayers, P. D., & Buschermohle, M. J. (2015). Perspectives on delineating management zones for variable rate irrigation. Computers and Electronics in Agriculture, 117, 154–167. https://doi.org/10.1016/j.compag.2015.06.019

Hornung, A., Khosla, R., Reich, R., Inman, D., & Westfall, D. G. (2006). Comparison of Site-Specific Management Zones: Soil-Color-Based and Yield-Based. Agronomy Journal, 98(2), Article 2. https://doi.org/10.2134/agronj2005.0240

ISPA. (2022). International Society of Precision Agriculture. Precision Ag Definition. International Society of Precision Agriculture. https://ispag.org/about/definition

Javadi, S. H., Guerrero, A., & Mouazen, A. M. (2022). Clustering and Smoothing Pipeline for Management Zone Delineation Using Proximal and Remote Sensing. Sensors (Basel, Switzerland), 22(2), 645. https://doi.org/10.3390/s22020645

Jaynes, D. B., Colvin, T. S., & Kaspar, T. C. (2005). Identifying potential soybean management zones from multi-year yield data. Computers and Electronics in Agriculture, 46(1), 309–327. https://doi.org/10.1016/j.compag.2004.11.011

Kissel, D., & Sonon, L. (2011). Soil Test Handbook for Georgia. The University of Georgia.

Lajili, A., Cambouris, A. N., Chokmani, K., Duchemin, M., Perron, I., Zebarth, B. J., Biswas, A., & Adamchuk, V. I. (2021). Analysis of Four Delineation Methods to Identify Potential Management Zones in a Commercial Potato Field in Eastern Canada. Agronomy, 11(3), 432. https://doi.org/10.3390/agronomy11030432

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. Computers & Geosciences, 34(9), 1044–1055. https://doi.org/10.1016/j.cageo.2007.07.010

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2022). cluster: Cluster Analysis Basics and Extensions. https://CRAN.R-project.org/package=cluster

Mallarino, A. (2001). Management Zones Soil Sampling: A Better Alternative to Grid and Soil Type Sampling? Proceedings of the Integrated Crop Management Conference, 11977198. https://doi.org/10.31274/icm-180809-717

Mallarino, A. P., & Wittry, D. J. (2004). Efficacy of Grid and Zone Soil Sampling Approaches for Site-Specific Assessment of Phosphorus, Potassium, pH, and Organic Matter. Precision Agriculture, 5(2), Article 2. https://doi.org/10.1023/B:PRAG.0000022358.24102.1b

Marcaida, M., Workman, K., & Ketterings, Q. M. (2025). Implication of soil grid sampling on lime, phosphorus, and potassium management of corn. Agronomy Journal, 117(3), e70074. https://doi.org/10.1002/agj2.70074

McFadden, J., Njuki, E., & Griffin, T. (2023). Precision Agriculture in the Digital Era: Recent Adoption on U.S. Farms.

Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F (2025). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. R package version 1.7-17, https://CRAN.R-project.org/package=e1071

Miao, Y., Mulla, D. J., & Robert, P. C. (2018). An integrated approach to site-specific management zone delineation. Frontiers of Agricultural Science and Engineering, 0(0), 0. https://doi.org/10.15302/J-FASE-2018230

Mitchell, C., & Huluka, G. (2012). Nutrient Recommendation Tables for Alabama Crops.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A. M. (2017). Delineation of Soil Management Zones for Variable-Rate Fertilization. In Advances in Agronomy (Vol. 143, pp. 175–245). Elsevier. https://doi.org/10.1016/bs.agron.2017.01.003

Oliver, M. A., & Webster, R. (1990). Kriging: A method of interpolation for geographical information systems. International Journal of Geographical Information Systems, 4(3), 313–332. https://doi.org/10.1080/02693799008941549

Pedersen, S. M., & Lind, K. M. (Eds.). (2017). Precision Agriculture: Technology and Economic Perspectives. Springer International Publishing. https://doi.org/10.1007/978-3-319-68715-5

Peralta, N. R., & Costa, J. L. (2013). Delineation of management zones with soil apparent electrical conductivity to improve nutrient management. Computers and Electronics in Agriculture, 99, 218–226. https://doi.org/10.1016/j.compag.2013.09.014

Pusch, M., Samuel-Rosa, A., Sergio Graziano Magalhães, P., & Rios Do Amaral, L. (2023). Covariates in sample planning optimization for digital soil fertility mapping in agricultural areas. Geoderma, 429, 116252. https://doi.org/10.1016/j.geoderma.2022.116252

Rojo Baio, F. H., Neves, D. C., & Teodoro, P. E. (2019). Soil Chemical Attributes, Soil Type, and Rainfall Effects on Normalized Difference Vegetation Index and Cotton Fiber Yield Variability. Agronomy Journal, 111(6), Article 6. https://doi.org/10.2134/agronj2019.05.0375

Sawyer, J. E. (1994). Concepts of Variable Rate Technology with Considerations for Fertilizer Application. Journal of Production Agriculture, 7(2), 195–201. https://doi.org/10.2134/jpa1994.0195

Schepers, A. R., Shanahan, J. F., Liebig, M. A., Schepers, J. S., Johnson, S. H., & Luchiari, A. (2004). Appropriateness of Management Zones for Characterizing Spatial Variability of Soil Properties and Irrigated Corn Yields across Years. Agronomy Journal, 96(1), Article 1. https://doi.org/10.2134/agronj2004.1950

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia, 126(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Shaner, D. L., Khosla, R., Brodahl, M. K., Buchleiter, G. W., & Farahani, H. J. (2008). How Well Does Zone Sampling Based on Soil Electrical Conductivity Maps Represent Soil Variability? Agronomy Journal, 100(5), Article 5. https://doi.org/10.2134/agronj2008.0060

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of Remote Sensing in Precision Agriculture: A Review. Remote Sensing, 12(19), Article 19. https://doi.org/10.3390/rs12193136

Soil Survey Staff. (2025). Web Soil Survey—Home. Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Alabama and Georgia. https://websoilsurvey.nrcs.usda.gov/app/

Song, X., Wang, J., Huang, W., Liu, L., Yan, G., & Pu, R. (2009). The delineation of agricultural management zones with high resolution remotely sensed data. Precision Agriculture, 10(6), 471–487. https://doi.org/10.1007/s11119-009-9108-2

Sudduth, K., Drummond, S. T., & Myers, D. (2012). Yield Editor 2.0: Software for Automated Removal of Yield Map Errors. In American Society of Agricultural and Biological Engineers Annual International Meeting 2012, ASABE 2012 (Vol. 4). https://doi.org/10.13031/2013.41893

Taylor, J. A., McBratney, A. B., & Whelan, B. M. (2007). Establishing Management Classes for Broadacre Agricultural Production. Agronomy Journal, 99(5), 1366–1376. https://doi.org/10.2134/agronj2007.0070

Thompson, A. N., Shaw, J. N., Mask, P. L., Touchton, J. T., & Rickman, D. (2004). Soil Sampling Techniques for Alabama, USA Grain Fields. Precision Agriculture, 5(4), 345–358. https://doi.org/10.1023/B:PRAG.0000040805.01967.b0

Tucker, M. W. (2024). Investigating the Effectiveness of Different Precision Soil Sampling Strategies for Site-Specific Nutrient Management in the Southeastern US [University of Georgia]. https://esploro.libs.uga.edu/esploro/outputs/graduate/INVESTIGATING-THE-EFFECTIVENESS-OF-DIFFERENT-PRECISION/9949666126302959

Valente, D. S. M., Pereira, G. W., de Queiroz, D. M., Zandonadi, R. S., Amaral, L. R. do, Bottega, E. L., Costa, M. M., de Freitas Coelho, A. L., & Grift, T. (2024). Accuracy of Various Sampling Techniques for Precision Agriculture: A Case Study in Brazil. Agriculture, 14(12), 2198. https://doi.org/10.3390/agriculture14122198

Veloso, F. R., Marques, D. J., De Melo, E. I., Bianchini, H. C., Maciel, G. M., & De Melo, A. C. (2023). Different soil textures can interfere with phosphorus availability and acid phosphatase activity in soybean. Soil and Tillage Research, 234, 105842. https://doi.org/10.1016/j.still.2023.105842

Veris Technologies. (2001). Frequently asked questions about soil electrical conductivity. Veris Technologies, Salina KS. https://www.veristech.com

Virk, S., Harris, G., & Lessl, J. (2024). Soil Sampling Grid Size Considerations for Site-Specific Nutrient Management. University of Georgia Extension.

Virk, S., Tucker, M., Harris, G., Smith, A., Levi, M., & Lessl, J. (2025). Efficacy and Economics of Different Soil Sampling Grid Sizes for Site-Specific Nutrient Management in Southeastern USA. Agronomy, 15(4), 903. https://doi.org/10.3390/agronomy15040903

Wakeel, A., Hassan, A., & Aziz, T. (2002). Effect of Different Potassium Levels and Soil Texture on Growth and Nutrient Uptake of Maize. 39.

Walton, J. C., Lambert, D. M., Roberts, R. K., Larson, J. A., English, B. C., Larkin, S. L., Martin, S. W., Marra, M. C., Paxton, K. W., & Reeves, J. M. (2008). Adoption and Abandonment of Precision Soil Sampling in Cotton Production. https://doi.org/10.22004/AG.ECON.46556

Wang, T., Sieverding, H., Elliot, M., & Cheye, S. (2023). Impact of High Fertilizer Prices and Farmers' Adaptation Strategies in the U.S. Midwest. Agriculture and Applied Economics Association, 38.

Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. Remote Sensing of Environment, 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

Wickham, H., Hester, J., & Bryan, J. (2024). readr: Read Rectangular Text Data. R package version 2.1.5. https://github.com/tidyverse/readr, https://readr.tidyverse.org

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77, 1–17. https://doi.org/10.18637/jss.v077.i01

Appendix A

Supplemental Information for Chapter Two

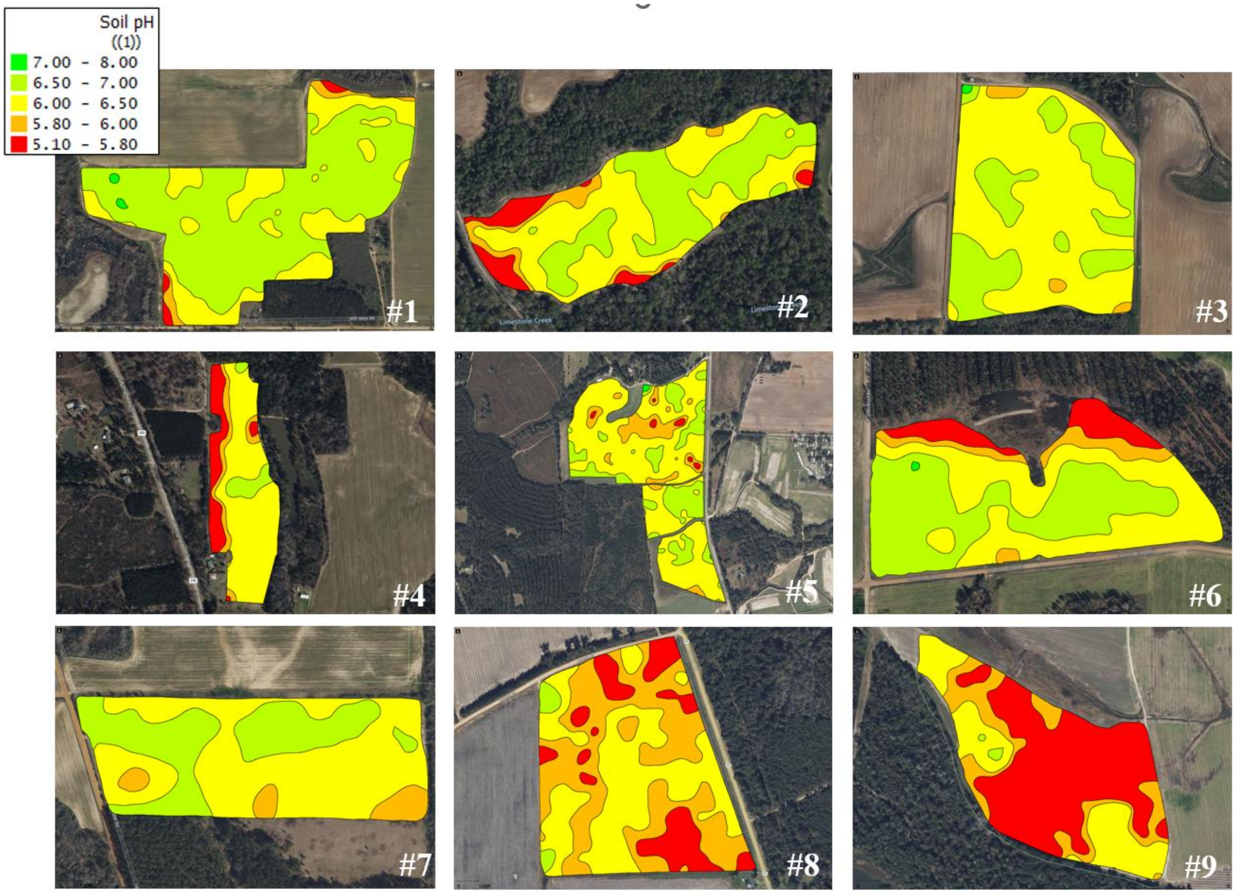

Figure A.1 Spatial maps demonstrating the soil pH variability in the fields (1 – 9) used in this study.

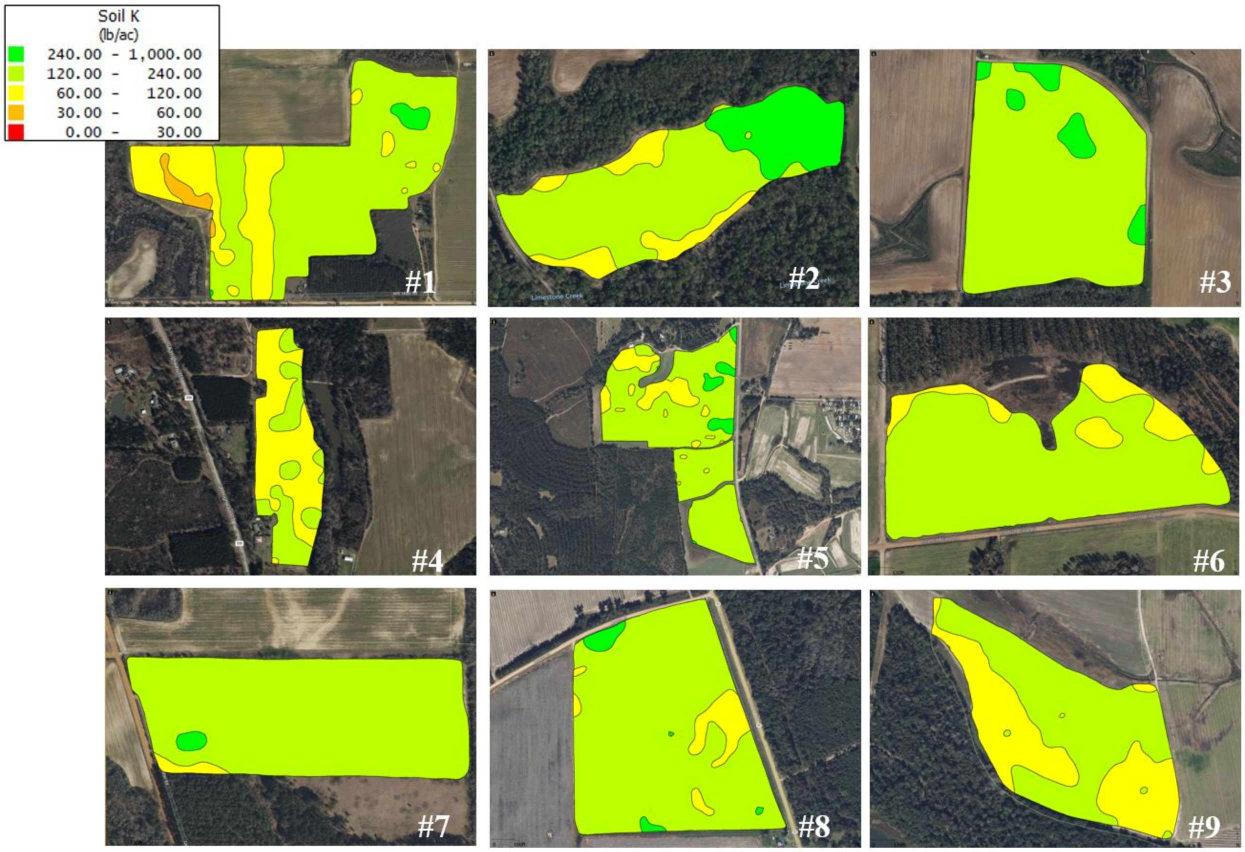

Figure A.2. Spatial maps demonstrating the spatial variability in soil K across the fields (1 - 9) used in this study.

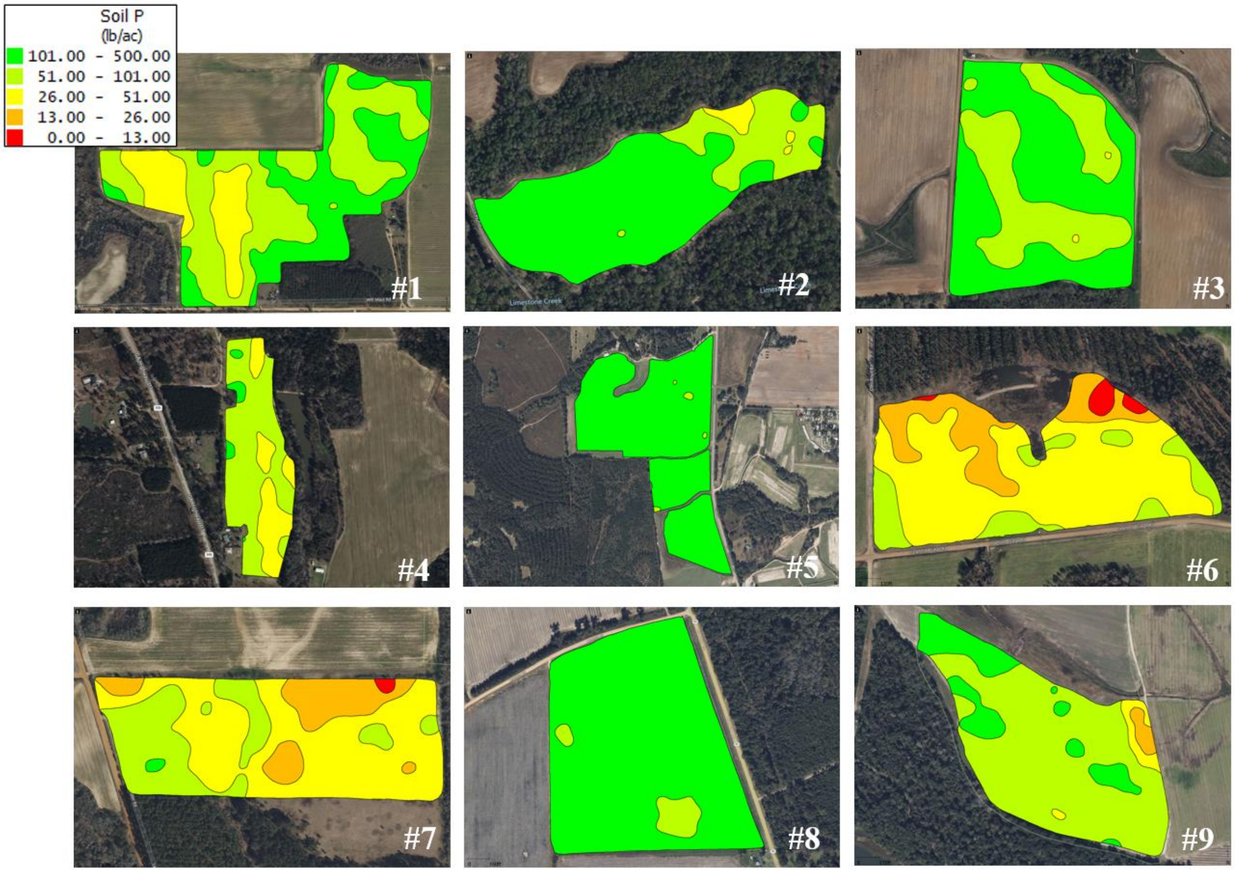

Figure A.3 Spatial maps demonstrating the soil P variability in the fields (1 - 9) used in this study.

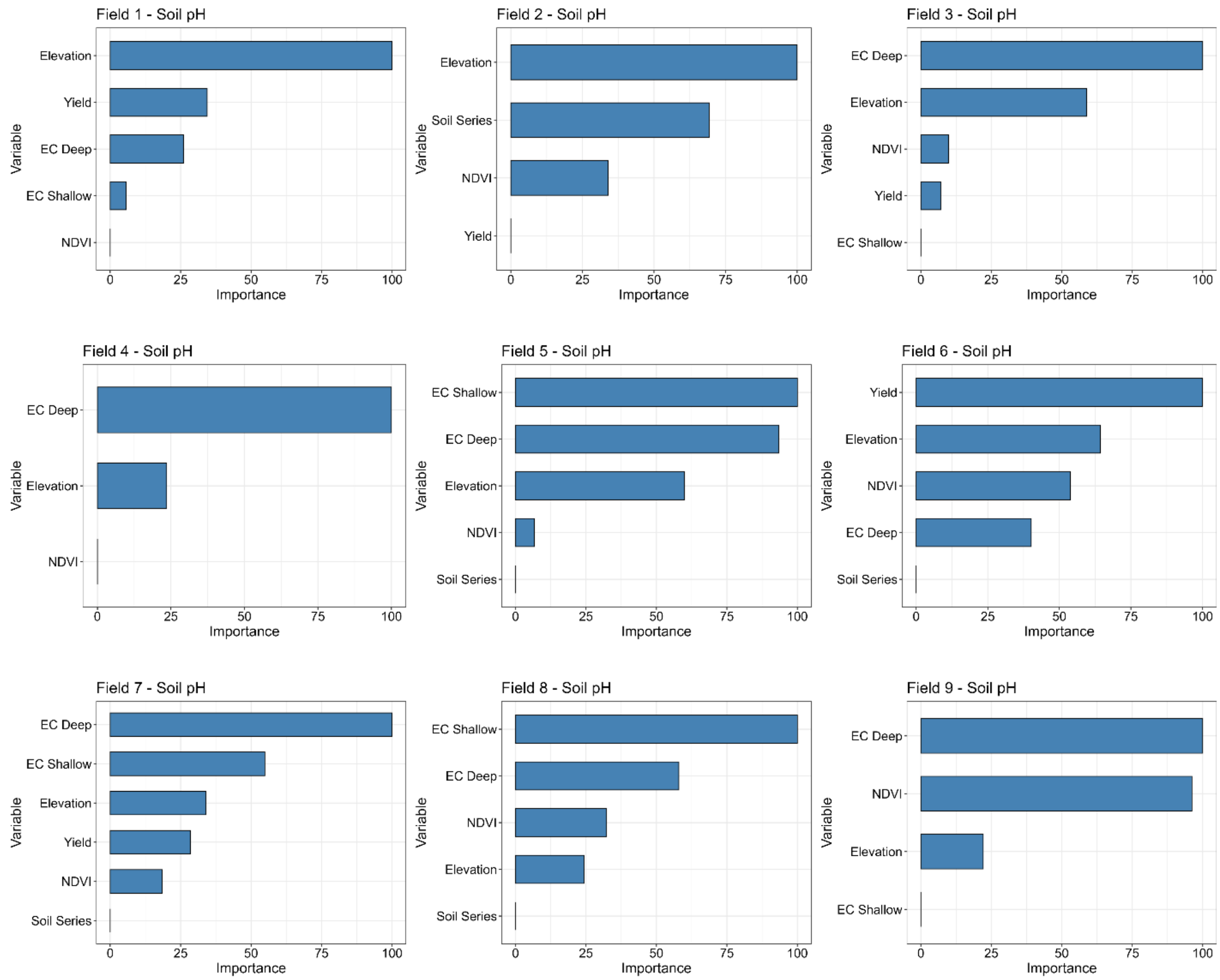

Figure A.4. Graphs showing variable importance for soil pH based on Random Forest modelling.

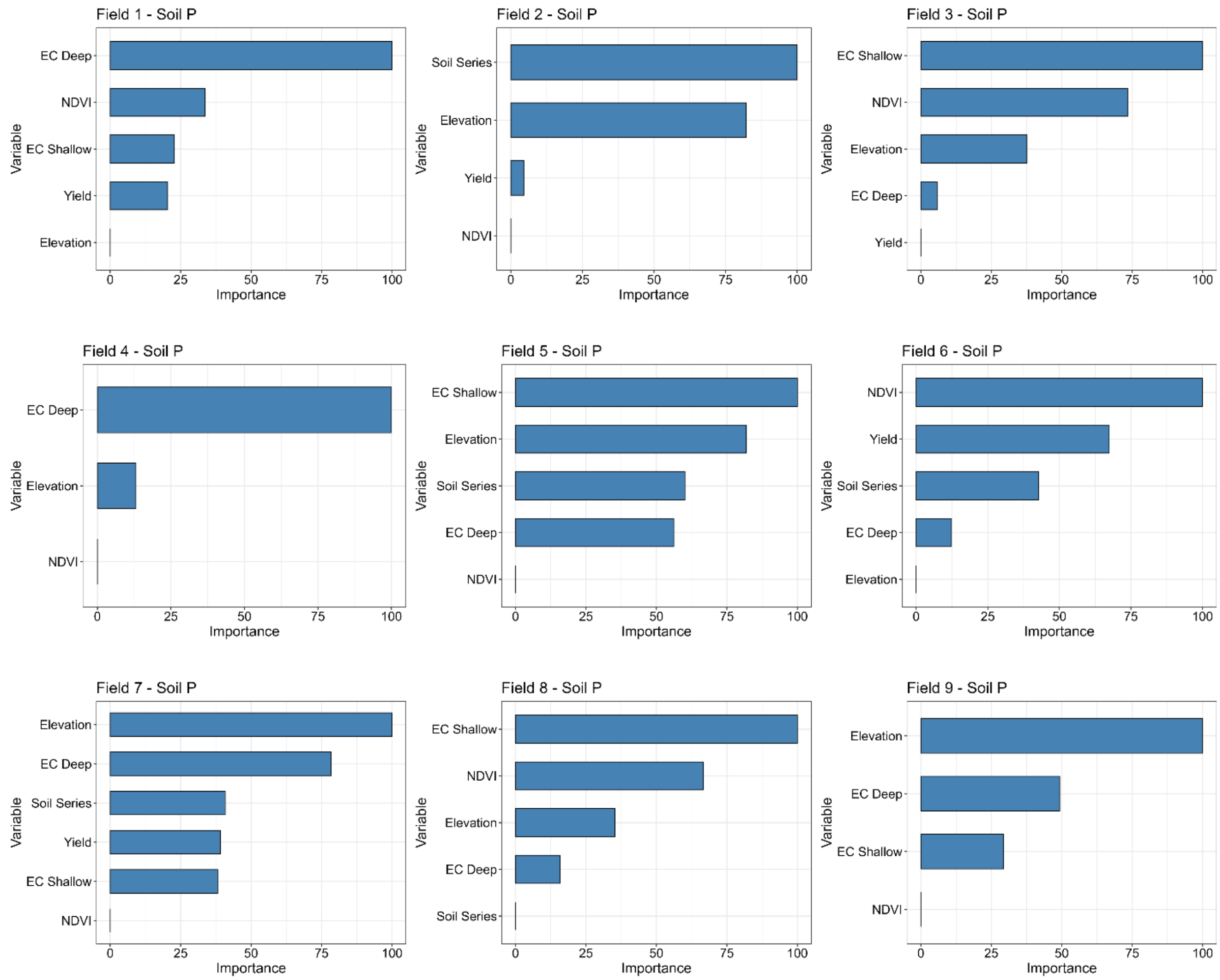

Figure A.5. Graphs showing variable importance for soil P based on Random Forest Modelling.

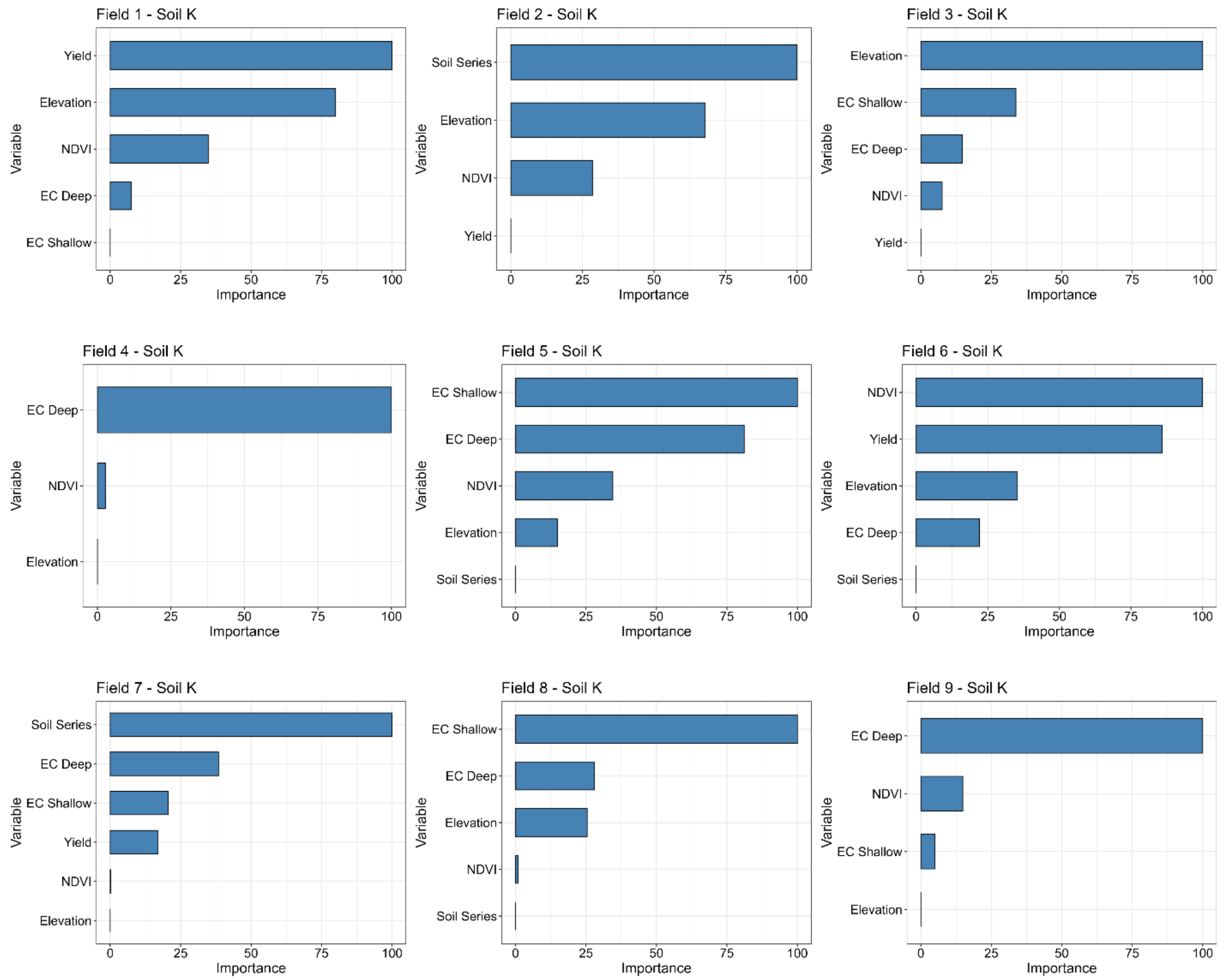

Figure A.6. Graphs showing variable importance for soil K based on Random Forest Modelling.

Appendix B

Supplemental Information for Chapter Three

Table B.1 Mean zone values for Field 1, k = 3.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	160 ab	118 a	6.4 b
	2	145 b	78 c	6.6 a
	3	173 a	98 b	6.6 a
S2	1	152 a	84 c	6.7 a
	2	161 a	114 a	6.4 b
	3	170 a	99 b	6.6 a
S3	1	151 a	123 a	6.5 b
	2	162 a	96 b	6.5 ab
	3	164 a	92 b	6.6 a
S4	1	168 a	97 a	6.6 a
	2	165 a	94 a	6.6 a
	3	135 b	108 a	6.3 b

Table B.2 Mean zone values for Field 1, k = 4.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	143 b	81 b	6.6 a
	2	147 b	127 a	6.3 b
	3	157 ab	81 b	6.6 a
	4	175 a	108 a	6.6 a
S2	1	179 a	111 a	6.6 a
	2	155 b	84 b	6.6 a
	3	158 ab	116 a	6.4 b
	4	156 b	85 b	6.6 a
S3	1	153 ab	113 ab	6.5 a
	2	169 a	95 bc	6.6a
	3	163 ab	90 c	6.6 a
	4	135 b	124 a	6.2 b
S4	1	128 b	78 c	6.6 a
	2	173 a	101 ab	6.6 a
	3	171 a	97 b	6.6 a
	4	133 b	124 a	6.2 b

Table B.3 Mean zone values for Field 2, k = 3.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	233 a	120 b	6.5 a
	2	223 a	120 b	6.4 a
	3	194 b	155 a	6.2 b
S2	1	242 a	114 b	6.4 a
	2	193 b	143 a	6.4 a
	3	234 a	119 b	6.4 a
S3	1	233 a	153 a	6.4 a
	2	232 a	117 b	6.4 a
	3	189 b	134 ab	6.4 a
S4	1	199 b	123 b	6.5 a
	2	238 a	121 b	6.4 a
	3	187 b	183 a	6.3 a

Table B.4 Mean zone values for Field 2, k = 5.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	225 b	119 bc	6.5 a
	2	176 c	169 a	6.2 b
	3	218 b	107 c	6.5 a
	4	226 b	136 b	6.3 ab
	5	300 a	124 bc	6.4 ab
S2	1	226 bc	114 b	6.4 ab
	2	187 d	164 a	6.3 b
	3	199 cd	120 b	6.5 a
	4	238 b	122 b	6.4 b
	5	300 a	124 ab	6.4 ab
S3	1	241 a	138 b	6.3 a
	2	256 a	143 ab	6.5 a
	3	193 ab	192 a	6.3 a
	4	234 a	117 b	6.4 a
	5	193 ab	121 b	6.4 a
S4	1	223 b	114 b	6.4 a
	2	187 c	183 a	6.3 a
	3	237 b	119 b	6.4 a
	4	195 c	129 b	6.4 a
	5	295 a	120 b	6.3 a

Table B.5 Mean zone values for Field 3, k = 3.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	174 a	44 a	6.5 a
	2	155 b	39 a	6.1 b
	3	198 a	36 a	6.5 a
S2	1	178 a	43 a	6.6 a
	2	200 a	34 a	6.6 a
	3	155 b	41 a	6.1 b
S3	1	174 a	45 a	6.5 a
	2	196 a	35 ab	6.6 a
	3	140 b	32 b	5.8 b
S4	1	140 b	34 b	5.9 b
	2	196 a	35 ab	6.6 a
	3	178 a	45 a	6.5 a

Table B.6 Mean zone values for Field 3, k = 5.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	176 b	46 a	6.5 a
	2	231 a	30 b	6.3 ab
	3	171 bc	44 ab	6.5 a
	4	181 ab	39 ab	6.5 a
	5	145 c	31 b	5.9 b
S2	1	152 b	39 a	6.1 b
	2	183 ab	36 a	6.7 a
	3	162 b	47 a	6.2 b
	4	231 a	30 a	6.3 ab
	5	179 b	43 a	6.6 a
S3	1	181 a	44 ab	6.6 a
	2	140 b	33 bc	5.9 c
	3	179 a	52 a	6.4 b
	4	100 b	16 c	5.5 d
	5	199 a	34 bc	6.5 ab
S4	1	204 a	32 ab	6.5 ab
	2	112 d	18 b	5.6 c
	3	181 ab	44 a	6.7 a
	4	172 bc	46 a	6.3 b
	5	150 c	47 a	6.2 b

Table B.7 Mean zone values for Field 4, k = 3.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	220 a	30 b	6.7 a
	2	185 b	51 a	6.4 b
	3	185 b	57 a	6.2 c
S2	1	172 b	44 b	6.3 c
	2	195 a	59 a	6.4 b
	3	211 a	26 c	6.6 a
S3	1	171 b	44 b	6.3 b
	2	194 a	58 a	6.4 b
	3	216 a	25 c	6.6 a
S4	1	172 b	55 a	6.3 b
	2	216 a	25 b	6.6 a
	3	193 a	53 a	6.4 b

Table B.8 Mean zone values for Field 4, k = 4.

MZ Delineation Strategy	Zone ID	Soil K (kg/ha)	Soil P (kg/ha)	Soil pH
S1	1	217 a	25 b	6.6 a
	2	202 ab	46 ab	6.6 a
	3	183 b	53 a	6.3 b
	4	185 ab	59 a	6.2 b
S2	1	211 a	23 c	6.6 a
	2	191 ab	66 a	6.4 bc
	3	203 a	45 b	6.4 ab
	4	170 b	46 b	6.2 c
S3	1	172 b	44 b	6.3 b
	2	206 a	22 c	6.6 a
	3	182 ab	72 a	6.4 ab
	4	199 a	57 a	6.4 ab
S4	1	216 a	25 c	6.6 a
	2	171 b	54 ab	6.2 b
	3	207 a	59 a	6.5 a
	4	170 b	44 b	6.3 b