EVALUATING THE FACTORS INFLUENCING FARMERS' CHOICES OF MAIZE-BASED CROPPING PATTERNS AND ASSESSING THE POTENTIAL OF DESIS HYPERSPECTRAL SATELLITE DATA TO DISCRIMINATE THE CROPPING PATTERNS.

CHARLYNNE JEPKOSGEI
June 2023

SUPERVISORS:
Prof.dr. A. (Andy) Nelson
Dr. R. (Roshanak) Darvishzadeh

EVALUATING THE FACTORS INFLUENCING FARMERS’ CHOICES OF MAIZE-BASED CROPPING PATTERNS AND ASSESSING THE POTENTIAL OF DESIS HYPERSPECTRAL SATELLITE DATA TO DISCRIMINATE THE CROPPING PATTERNS.

CHARLYNNE JEPKOSGEI
Enschede, The Netherlands, June 2023

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of requirements of Master of Science in Spatial Engineering.

SUPERVISORS:
Prof.dr. A. (Andy) Nelson
Dr. R. (Roshanak) Darvishzadeh
Mbali Mahlayeye (PhD) - (Advisor)

THESIS ASSESSMENT BOARD:
Dr.IR. (Anton). Vrieling (Chair)
Professor Abel Ramoelo (External Examiner, Director Centre for Environmental Studies, University of Pretoria)

ABSTRACT

Cropping patterns are defined as the annual sequence and spatial arrangement of crops on a piece of land, for example monocropping and intercropping patterns. Information about the location, extent, and types of cropping patterns is crucial for accurately measuring crop production and land use intensity for food security assessment. However, underlying factors that influence farmers' choices about which cropping pattern to adopt are still unclear in most regions in the world. Technology such as remote sensing has enabled the mapping of cropping patterns, though this is challenging in regions dominated by small-holder agriculture in regions like Sub-Saharan Africa due to a variety of limitations such as small field size, highly fragmented landscapes, and the nature of the cropping patterns. This study first explores local knowledge from field farmers' survey responses by evaluating the factors influencing the choice of monocropped and intercropped maize patterns in Busia, Kenya. It then highlights and discusses six factors, including size of the fields, household needs, availability of resources, farmers' experience/preference, market demand, pest control/plant symbiosis in comparison to existing literature. Further, the study assesses the discrimination of maize cropping patterns using DESIS hyperspectral satellite data to characterize monocropped and intercropped maize fields. We extracted reflectance of the fields of both cropping patterns based on the field boundary data. Statistical tests identified the bands that showed significant spectral differences between monocropped and intercropped maize fields. A Random Forest (RF) classifier was used for feature selection to identify the best subset of features (bands) that would further be used for classification. The results from the statistical tests indicated a statistical difference in the spectral signatures of the two cropping patterns. As such, 110 significant bands were identified in the visible, red-edge and near-infrared (NIR) spectral regions that were the most sensitive to discriminating the cropping patterns. From feature selection, five bands dominated in the red edge and NIR (752.2nm, 767.5nm, 775.2nm, 783nm, 814.2nm) were further selected for classification. Those bands were used in a RF classifier, obtaining an overall accuracy (OA) of 74% with a producer accuracy (PA) of 71% for monocropped maize fields and 80% for intercropped maize fields and user accuracy (UA) of 91% for monocropped fields and 50% for intercropped fields. An F1 score of 80% for monocropped and 62% for intercropped maize fields was obtained. A kappa coefficient of 0.43 was attained, indicating the complexity of discriminating and classifying the maize-based cropping patterns. The results of this study show that there is potential discrimination of the maize cropping patterns using hyperspectral remote sensing, but it can be quite challenging especially in the late development stage of maize. Hence, remote sensing images need to be obtained during the early part of the growing season before the maize canopy obscures the smaller intercropped crops. The exploratory nature of this research opens more avenues for future research into cropping patterns discrimination in small-holder agriculture and further suggests that a combination of narrative perspectives from farmers and the use of remote sensing technologies is required for an in-depth understanding and addressing current food security challenges.

Key words: Farmers' interview; cropping patterns; hyperspectral data; Mann-Whitney *U* test; feature selection, random forest.

ACKNOWLEDGEMENTS

To my supervisors, Andy Nelson and Roshanak Darvishzadeh, thank you for accepting my proposal for this topic and for guiding me diligently through this MSc research. I appreciate the constructive remarks, suggestions and encouragements and assuring that I give nothing but the best. I have learnt a lot from both of you throughout the supervision. Mbali Mahlayeye, thank you for the advice and guidance on this research as it stemmed out from your interesting PhD work. Kelvin Mlawa, thank you for being such as amazing teammate during fieldwork and throughout the MSc journey. You have all been very instrumental and supportive.

I thank Anton Vrieling for being very supportive and instrumental as a chair of the thesis board as well as my internship supervisor. I also extend my valuable thank you to Raymond Nijmeijer for being a great mentor throughout my MSc programme.

I would like to also extend my gratitude to Orange Knowledge Programme for the scholarship opportunity that has enabled me to pursue this master's and maintain my stay here in the Netherlands. I am grateful to ITC Faculty as well for being financially supportive in my education and research work.

Special appreciation goes to all the farmers who were willing to participate in the field data and surveys that enabled me to do this research.

To my friends and mentors who have really been supportive throughout this journey with encouraging words, thank you.

Exceedingly, I am grateful for my family; mum, dad, and my beloved sisters for the support you have given me through my time away. I would have loved to repay in huge lumpsums someday, my love for you is deep and enormous.

Lakeisha my daughter, thank you for being the light of my heart. I have spent my time away and with my laptop more than you but your innocent patience and love you have upheld for all this while are indispensable. You have been my strength and motivation. I love you more than anything else.

Above all, I am grateful to God for the gift of life, good health, great opportunities and support that I have experienced throughout my MSc.

Charlynne Jepkosgei,
Enschede,
The Netherlands, June 2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Food insecurity: A wicked problem

Food security exists when people have sustainable physical and economic access to adequate, safe, nutritious, and socially acceptable food for a healthy and productive life (FAO, 2017). Determinants of food security include availability, access, stability, and utilization (FAO, 2017; Peng & Berry, 2018). This study focuses on the food availability dimension of food security. Food availability is associated with geographical and agricultural determinants, such as domestic production, import capacity, food stocks, water availability, agriculture land, population density and distribution (Peng & Berry, 2018). Moreover, a study by Kane (2014) states that this food security dimension is influenced on a local and global level by three important factors which include "access and availability of food in local environments, effects of the changing climate on agriculture and natural resources, and active participation in planning, developing and managing effective strategies to optimize and sustain food production with the available existing land."

Over one-third of Africa's population is considered severely food insecure (FAO, 2021). In Kenya, a combination of factors, including poverty, climate change, crop and livestock diseases, ineffective policies, and market-related bottlenecks, combine to aggravate the problem of food insecurity in many households. About 90% of primary food production from agriculture feeds a population of 53 million people (USAID, 2021), yet only about 20% of Kenyan land is suitable for farming. A combination of factors such as climate variability, poor soils, lack of inputs and technology, and crop pests and diseases depress yields and have complex and compounding effects on food systems, negatively impacting the country's food security (USAID, 2021). Also, as the population increases, it results in increased food consumption while resources such as agricultural land and water required for production remain limited. To account for these challenges, there is a need for the expansion of agricultural land for crop production which may cause other challenges such as ecosystem degradation, and conversion of forests, savanna, and wetlands to increase which in the long run, contributes to greenhouse gas (GHG) emissions that in turn contribute to climate change. In addition, a combination of other factors including high and persistent levels of inequality, economic slowdowns that were recently exacerbated by the COVID-19 pandemic in 2020 and recent political instability, combine to aggravate the problem of food insecurity.

One of the most important strategies for improving food security is to adopt sustainable agricultural management practices, which in this context means sustainable use of existing land to improve food production. And one of these management practices is intercropping. Intercropping is an agronomic practice of growing two or more crops simultaneously on a piece of land (Asseng et al., 2014, Mthembu et al., 2019) while utilizing existing resources such as land and water for food production (Glaze-Corcoran et al., 2020). Several other studies (Maitra et al., 2020; Mthembu et al., 2019; Morris, 2017; Matusso et al., 2012; Rusinamhodzi et al., 2012) have reported the benefits of intercropping including but not limited to; a "backup" crop should one of the two crops fail, restoration of soil fertility, better management and control of pests, diseases and weeds, water conservation in soils and enhanced carbon sequestration and proper utilization of resources such as rainfall. In addition, in areas that are prone to extreme weather conditions especially with the current climate variability due to climate change, intercropping provides a greater "insurance" against the risk of crop failure (Schmutz, 2020). Also, since different crops are harvested from the same piece of land in intercropping pattern practices, for example, legumes are planted with cereals, families can enjoy better nutritional balance (Rusinamhodzi et al., 2012).

On the other hand, monocropping is a practice of planting and growing one type of crop on the same plot of land, year after year (Robbins, 2022). Monocropping is advantageous because it requires knowledge of just one crop and is relatively easy to manage. However, despite this economic advantage, the monocropping agricultural practice has posed several challenges for small holder farmers especially in sub–Saharan Africa. A study by Mthembu et al. (2019) and Robbins (2022) highlight the challenges posed by monocropping agricultural system such as crop failure and a decline in soil fertility. This is exacerbated by frequent pest, diseases, and weeds, resulting in reduced crop productivity and quality over time, reducing viable livelihood and households' food security (Robbins, 2022). However, despite the disadvantages of monocropping systems, intercropping also has its limitations, a major one being a potential yield reduction in the sense the different crops compete for resources as compared to a monocrop cropping pattern. This is discussed in depth by Mthembu et al., (2019). It can be seen as a trade-off between productivity and sustainability.

Sustainable Development Goal 2 (Target 2.4) states that "there is need to ensure sustainable food production systems and the implementation of resilient agricultural practices that increase productivity and production while maintaining a balanced ecosystem. These sustainable agricultural practices should strengthen the capacity for adaptation to climate change, extreme weather, drought, flooding, and other disasters." Despite intercropping pattern practices being labour and resource intensive, given its numerous advantages over monocropping practices, it can be deduced that intercropping practices can be a viable strategy for long-term food production and sustainability in many contexts. These land-use practices can improve agricultural resilience to climate change while reducing environmental impacts (Glaze-Corcoran et al., 2020; Asseng et al., 2014; Bégué et al., 2018). Interestingly, farmers in Kenya have been practising intercropping for a long time. For example, in areas where maize is prominently cultivated, other crops like beans, peas, potatoes, millet, cassava and groundnuts are introduced. However, where farmers practice intercropping and the factors that influence farmers' choice of this cropping pattern are unclear. Therefore, exploring methods to identify intercropping patterns and understanding the influencing factors from a farmers' perspective are crucial in designing targeted interventions to boost food security. This study explores this gap based on farmers' response from a field survey and uses DESIS hyperspectral satellite remote sensing data to capture the characteristics of these cropping patterns.

1.2. Need for enhanced agricultural management monitoring

Currently, available information about agricultural land use management practices at various spatial levels typically relies on conventional methods of information acquisition such as farmers interviews, land use surveys, and crop statistical data which are labour-intensive, time consuming and hence tedious and expensive to obtain (Ibrahim et al., 2021). In addition, data collected in such conventional ways may soon become outdated if not collected regularly. Although these data may contain rich information, there is a lack of detailed mapped data that provides information about specific land use management practices, particularly in rapidly changing environments (Khan et al., 2010). For instance, agricultural production in small-holder agriculture that is practised in dynamic and heterogenous landscapes has not been fully researched and mapped, making it a challenge to be characterized on a seasonal basis (Mthembu et al., 2019; Rusinamhodzi et al., 2012; Yonah et al., 2018).

In recent years with the current advancement in technology, robust and cost-effective strategies have been developed, which contribute to knowledge and understanding of food insecurity. These strategies include using satellite remote sensing and its related spatial analytical techniques to "examine local food environments, assess changes in land use and land cover, identifying areas of importance in specific regions to determine the relationships between biophysical and socioeconomic attributes of the food system and the linkages between sustainability, food security and climate change" (FAO, 2018). However, detailed information and understanding of cropping patterns, especially in small-holder farming systems, are lacking in many countries in Africa, Asia, and Latin America (Bégué et al., 2018; Ibrahim et al., 2021; Mahlayeye et al., 2022).

Since governments, NGOs, civil society organizations, development agencies, crop insurance agencies and even private sector companies focus on supporting small-holder farmers in farming more sustainably, understanding farming practices requires detailed descriptions of what and where crops are grown (Khan et al., 2010). Accurate, updated and geographically explicit information can be leveraged from advanced remote sensing technologies to identify which cropping patterns are practised where and when. In a food security context, identifying and characterizing these cropping patterns provides valuable information to assess changes in land management practices and provide a better estimation of food production. In addition, assessing the changes in land management practices plays a critical role in supporting and understanding current and future sustainability policies to reform the agricultural sector towards sustainability of food systems in the wake of climate change and increasing pest and disease risks to food production. In addition, the improvement of policies to support such good agricultural practices, might benefit the social, ecologic, and economic development of small-scale farmers, thereby reducing food insecurity (SDG 2) and alleviating poverty (SDG 1).

1.3. Remote sensing approaches to characterize cropping patterns

Over the past 50 years, numerous remote sensing-based methods and datasets have been developed and successfully applied in the field of agriculture. The use of multi-temporal optical imagery has been explored over the decades mostly for crop type mapping. However, few studies have utilized satellite imagery to map cropping patterns. Extensive review of remote sensing for mapping crops and periodic cropping practices and patterns have been carried out by Bégué et al., (2018) and Mahlayeye et al., (2022). The two reviews provide an overview of several scientific studies that have used multispectral broadband satellite data to map and classify crop varieties and cropping patterns by utilizing discrete bands and/or vegetation indices. Both reviews indicate the challenges of mapping or discriminating intercropping patterns from monocropping patterns. Bégué et al. (2018) reported that use of different remote sensing technologies such as the new Sentinel constellations, hyperspectral data and/or sensor data combinations can be explored to improve the mapping of cropping patterns in heterogenous landscapes. Mahlayeye et al. (2022) also acknowledged that the increased availability of different sensors provides an opportunity to advance the mapping of intercropping fields, especially in complex landscapes, such as those found in Sub-Saharan Africa.

A few studies have used very high spatial resolution data to discriminate maize cropping patterns in Kenyan landscapes. The study by Richard et al. (2017) successfully mapped maize cropping systems (mono- and intercrop) in Kenya using RapidEye bi-temporal data with an overall classification accuracy of 93% and class accuracies for the two cropping systems above 85%. The study concluded that high spatial resolution and multitemporal data is required to explicitly map or discriminate different cropping patterns/systems especially in a highly fragmented and heterogeneous landscape. Another study explored the combination of time series from multispectral data from Sentinel 2 and Synthetic Aperture Radar (SAR) from Sentinel 1 (Rebecca et al., 2020) to classify cropping patterns in the dynamic and heterogenous landscape of three sub-counties in Kenya. However, despite the high spatial resolution of these multispectral sensors, the reports still indicate a high misclassification rate because of the high inter-class variability caused by the different vegetation compositions. In addition to exploiting temporal and spatial resolution data, there is the opportunity to look at higher spectral resolution data, hyperspectral data, to improve the discrimination accuracy between different cropping patterns, that is, intercropping and monocropping (Mahlayeye et al., 2022). This research identified and explored this niche.

Hyperspectral sensors provide high spectral resolution, and studies have shown that the narrow spectral bands from these sensors can be used to distinguish crop varieties for a wide range of crops as well as different land use patterns (Agilandeeswari et al., 2022; Goetz, 2009; 2013; Marshall & Thenkabail, 2015). Studies by Bégué et al. (2018), Kot et al. (2017), Mariotto et al. (2013), Marshall et al. (2022) and Teke et al. (2013) reported that the level of spectral detail afforded by hyperspectral data enhances vegetation characteristics which are difficult to achieve with multispectral data. Other studies have shown that crop species can be discriminated at the leaf and canopy level based on their spectral reflectances, showing the improvement of hyperspectral imagery in crop discrimination and mapping of vegetation species and communities (Dian et al., 2009; Kumar et al., 2019; Schmidt & Skidmore, 2003; Shafri et al., 2011; Sobhan, 2007; Vaiphasa et al., 2005). Lu et al., (2020) and Teke et al., (2013) discussed a more detailed review of recent advances in hyperspectral remote sensing and its adoption in different agricultural applications. Despite these significant advances, hyperspectral remote sensing data has not been explored in agricultural applications to characterize or map cropping patterns.

1.4. Hyperspectral satellite data in crop discrimination

Previous studies that have utilized hyperspectral data show the possibility to discriminate crop and vegetation species but also state that the narrow spectral bands from hyperspectral data contain redundant information, which makes the computation difficult for discrimination and classification (Dian et al., 2009; Mariotto et al., 2013; Prospere et al., 2014; Shafri et al., 2011; Sobhan, 2007). Various univariate and multivariate algorithms have been proposed to reduce the redundancy and dimensionality of hyperspectral data. Studies have focused on approaches such as principal component analysis, partial least square regression, artificial neutral network, and statistical approaches such as the Mann-Whitney *U* test and ANOVA to understand the statistical differences in hyperspectral bands (Darvishzadeh, 2008; Adam & Mutanga, 2009; Schmidt & Skidmore, 2003; Shafri et al., 2011; Vaiphasa et al., 2005).

When it comes to classification, the high spectral dimensionality in hyperspectral data introduces multicollinearity into the data and reduces the accuracy of the classification algorithms (Adam et al., 2012). To address this, methods that employ feature selection algorithms as part of the evaluation process have been proposed before classification is performed (Thenkabail et al., 2004; Vaiphasa et al., 2005). These include wrapper feature selection algorithms that use the classification algorithm as part of the assessment process to look for the optimal subset of bands (Kohavi and John 1997; Kavzoglu and Mather 2002; Prospere et al., 2014; Sobhan, 2007) or the filter approach algorithms which analyses subsets of bands using the training data (Dian et al., 2009; Kumar et al., 2019; Schmidt & Skidmore, 2003).

The study by Adam et al. (2012) and Prospere et al. (2014) described different classification methods and indicated that Random forest (RF) tree based models are more robust compared to other classification methods - such as support vector machine (SVM), linear discriminant analysis (LDA), neural networks (ANN) and classification and regression trees (CART). RF entail building multiple classification trees from a selection of training samples data and identifying the best bands that contribute the most to the classification (Liaw & Wiener, 2002). More importantly, in hyperspectral images where the pixels in regions composed of a single class (e.g., vegetation species) can contain very different spectral signatures (Borsoi et al., 2021), RF is considered robust as the classifier minimizes error from a single decision tree by choosing random samples, creating several decision trees, and using a majority vote to reach a final decision. Several other studies have demonstrated that RF can be successfully used for feature selection as well as for classification (Díaz-Uriarte & Alvarez de Andrés, 2006; Granitto et al., 2006; Sabat-Tomala et al., 2020), though only a few remote sensing studies have utilized RF for feature selection and classification when hyperspectral satellite data is used (Adam et al., 2012; Prospere et al., 2014). This study considers these research gaps and aims to address them.

1.5. Problem statement

As crop production faces increasing challenges due to climate change, ecosystem degradation, and population pressure, the threat of an imbalance in food supply and demand will continue to rise both locally, nationwide, and worldwide. Intensive agricultural practices such as intercropping practices need to be adopted to increase food security for the growing population while utilizing existing/available resources. However, there is a lack of consistent, sufficient, and geographically coherent data on cropping patterns at local, regional, national, and global scales. In addition, the factors that influence farmers' choices on different cropping patterns are often location specific and need to be understood before further adoption of intercropping can be encouraged. Few farmers surveys have been conducted to understand the factors that influence the choice of crop types and crop patterns in sub–Saharan Africa. Moreover, information and specific characterization and mapping of these cropping patterns are missing from local, regional, and even national levels.

Thus, considering these research gaps, from understanding the choices of farmers in adopting different cropping patterns, in this case maize cropping patterns and that no research has utilized DESIS hyperspectral data to characterize these cropping patterns, the aim of this research is to explore these gaps using the data collected from field in Busia County, Kenya.

1.6. Research objectives

1.6.1. Main research aim

The main aim is to evaluate the factors influencing farmer's choice of maize-based cropping patterns in Busia County, Kenya and then determine the usefulness of DESIS hyperspectral satellite data to discriminate these cropping patterns.

1.6.2. Specific research objectives

To achieve the overall aim, the following specific objectives will be undertaken:

i. To evaluate farmers' survey data and characterize the factors influencing farmers' choice of cropping patterns.
ii. To analyze spectral differences between maize based cropping patterns using spectral signatures obtained from DESIS hyperspectral satellite data.
iii. To identify the optimal narrow bands that best discriminate between cropping patterns and use them for cropping pattern classification.

1.6.3. Research questions

i. What are the factors that influence farmers' choice of a cropping pattern?
ii. Which spectral bands/regions from the spectral signature of monocropped and intercropped maize are statistically different?
iii. What are the optimal bands that can discriminate cropping patterns and how robust is the resulting classification?

1.6.4. Hypothesis

i. Ho: There are no factors that influence farmers choices of cropping patterns.
   Ha: Physical and socio-economic factors influence farmers choices of cropping patterns.
ii. Ho: There is no statistical difference in spectral bands in hyperspectral data that can be effectively used to discriminate maize cropping patterns.
   Ha: There is statistical difference in the spectral bands in the red edge and NIR region of hyperspectral data that can be effectively used to discriminate maize cropping patterns.
iii. Ho: All wavelengths' regions are optimal are for maize cropping pattern discrimination and classification and the classification accuracy is high.
   Ha: The red edge and NIR region give the optimal bands for cropping pattern discrimination and classification and the classification accuracy is higher.

1.7. Expected research output.

The research mainly focuses on two core knowledge areas in Spatial Engineering: spatial planning and governance (SPG) and spatial information science (SIS). The SPG expected research outputs are based on farmers' responses to a survey, investigating the factors that affect farmers' choices for maize based cropping patterns. The SIS knowledge focuses on the remote sensing analysis. The objectives are based on outlining the factors that farmers perceive to influence their decisions on cropping patterns and identifying a subset of optimum bands that best discriminates maize cropping patterns using hyperspectral image data. The subset of bands identified will be further used for classification using an RF classifier. Ultimately, based on the results, this research will derive recommendations for further research on integrating the SPG and SIS in reducing food insecurity wickedness.

2. STUDY AREA AND DATA

This section provides the description of the study area, data and the software used during the research.

2.1. Study area

The study area is within Busia County, situated in the western region of Kenya (Figure 1). The study area's central coordinates are 0° 27' 38.77" N and 34° 06' 41.26" E, bordering Uganda to the west. It covers Teso North, Teso South, Nambale and part of Matayos sub-counties in Busia County. According to the County Government (CGOK) integrated Plan (2018-2022) (Busia, 2018), the county spans about 1700-kilometre square making it one of the smallest counties in Kenya. The county is bordered by Bungoma County to the north, Kakamega County to the east, Siaya to the southeast and Lake Victoria to the southwestern part (Busia, 2018).

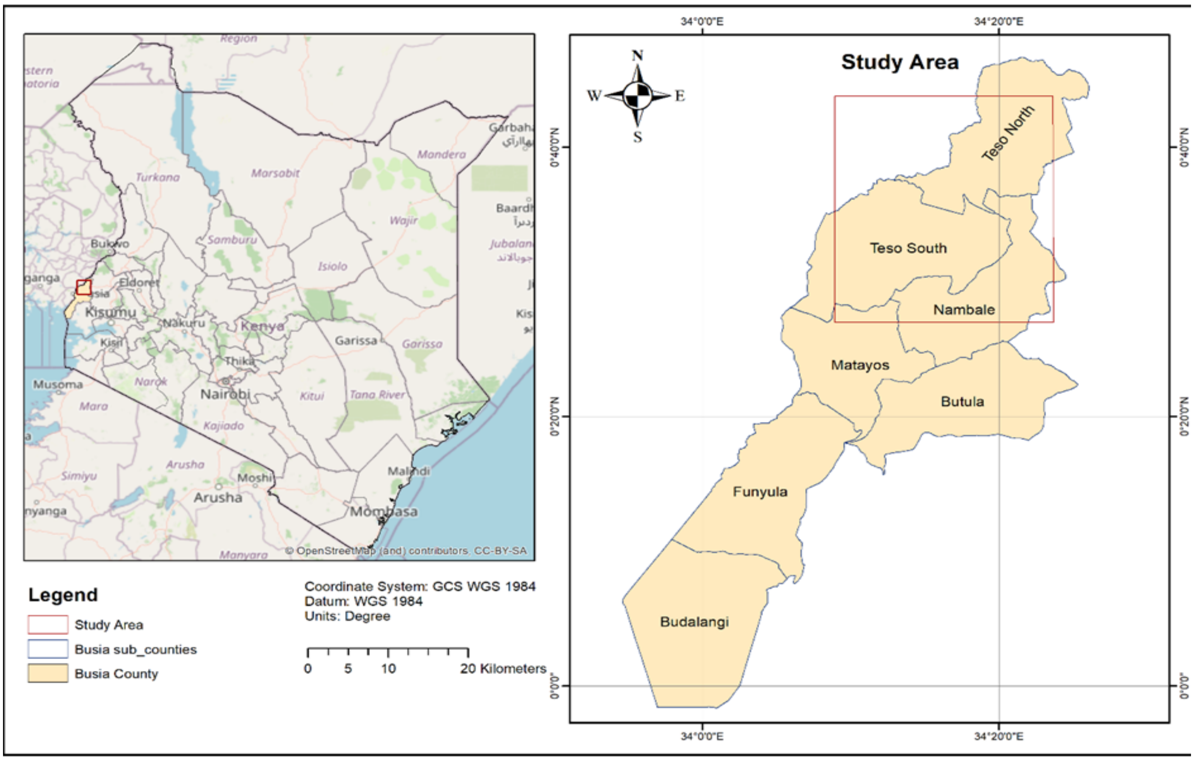

Figure 1: Study area and location of field boundaries used for the study.

Lake Victoria greatly affects and influences the county's climatic conditions. The annual rainfall is between 760 mm and 2000mm. There are two rainy seasons, long rains from March to May and short rains from September to December (Busia, 2018). The annual mean temperature ranges between 26 ºC and 30ºC, and the humidity of the air is relatively high due to the proximity to the lake. Figure 2 shows the average monthly rainfall in Busia in 2021 and 2022 (climateSERV). According to the County Government (CGOK) integrated Plan (Busia, 2018), the main economic activity is agriculture, which is practised extensively throughout the two rainy seasons. The dominant food grown is maize, followed by cassava, sorghum and millet, cotton, tobacco, and sugarcane (Busia, 2018). Maize farms, however, are predominantly planted with other crops like beans, soya and groundnuts, indicating intercropping is widely practised, making it a region of interest to study. The average farm size in the county is approximately 0.7ha, with the small average farms ranging from 0.2 - 0.3ha (Hickey et al., 2012).

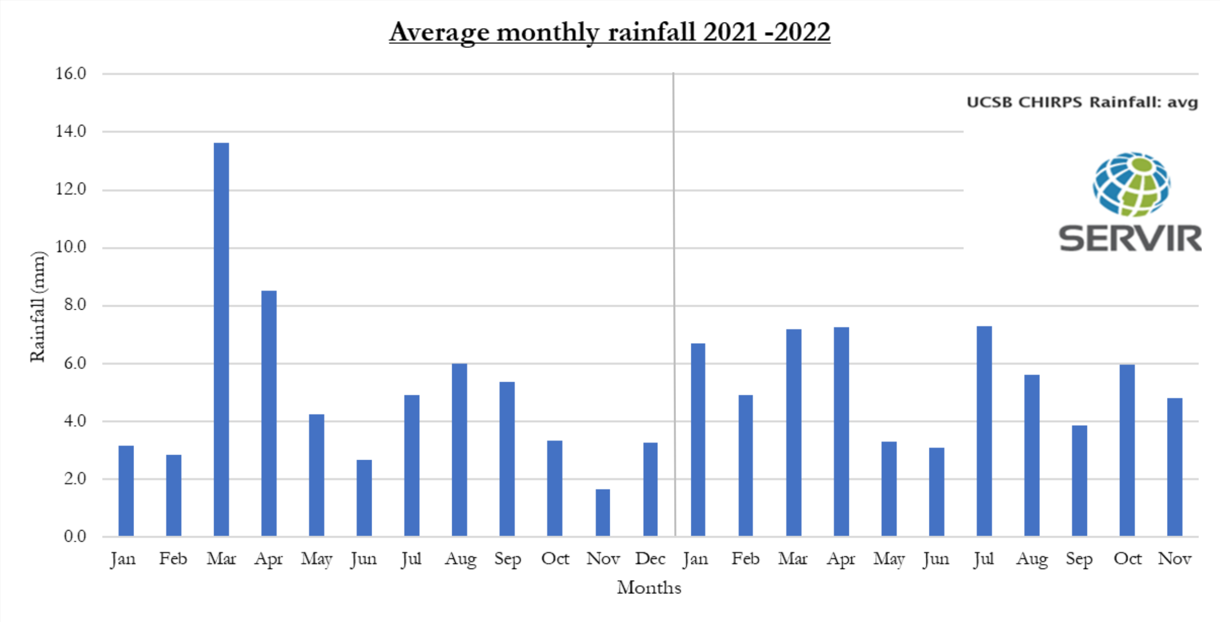

Figure 2: Average monthly rainfall in Busia County in 2021 and 2022 (source: climateserv.servirglobal.net).

2.2. Data

The research utilized hyperspectral satellite data and field data based on farmers' survey and field boundary data collection. The following subsections describe the datasets.

2.2.1. DESIS Hyperspectral data

DESIS which stands for DLR Earth Sensing Imaging Spectrometer, is a hyperspectral instrument from the German Aerospace Center (DLR, 2019). DESIS hyperspectral data covers the visible and near Infrared regions of the electromagnetic spectrum ranging from 400 – 1000nm with 235 narrowband channels each of 2.5nm width. The data is obtained with a pixel resolution (ground sample distance) of 30m from a 400km orbit altitude (Alonso et al., 2019; (DLR, 2019). Table 1 shows the characteristics of the DESIS sensor as reviewed by Eckardt et al. (2015). We obtained one DESIS image captured on 7th of June 2021 at level 2A (L2A) surface reflectance in GeoTIFF format, stored in integer radiance values. This image was one year prior to the fieldwork, and thus our field surveys relied on farmer recall for cropping practices and crop calendar information in the previous year. In addition, we later obtained PRISMA hyperspectral data which was acquired on 9th July 2022 during the field collection. However, this 2022 hyperspectral image could not be used for this research because it was tasked late in the season when most fields(farms) were already harvested.

Table 1: Characteristics of the DESIS instrument (sensor) (source:Eckardt et al. (2015)).

	DESIS
Telescope	2.8 / 320 mm
FOV	4.4°
IFOV	0.004°
GSD	30m@400km
BRDF angle	+/- 40 ° in flight direction
Spectral range	400-1000nm
Spectral sampling	2.55 nm
Spectral channels	235
Spatial channels	1024
SNR	> 150 (@2.5nm)
Radiometric linearity	>98% (10%-90% FWC)
Instrument MTF @Nyquist without smearing	Between 20%-25%
In Flight Calibration	Internal LED field and dark / DSNU/PRNU/Geometry/ Linearity

2.2.2. Field data

This study utilized a rich set of data collected from the field in July-August 2022 in Busia, Kenya. Four sub-counties were visited on different dates and farmers from the different sub-counties were interviewed. Farmers interviews were based on crop calendar information, challenges, and the factors influencing their choices for cropping pattern were recorded in a questionnaire form (Appendix A) in the different agroclimatic zones in Busia. The crop calendar information was based on the crops grown during the first and second seasons of 2021 and the first season of 2022. In addition, field boundaries were collected as polygons. The sampling scheme and fieldwork are described in detail in Chapter 3, and the descriptive analysis of the farmer's interviews is further explained in Chapter 4.

Figure 3: Farmers' interviews during the fieldwork in Busia.

2.3. Ethical consideration and data management

The ethical aspects of this research were assessed by the ITC GEO Ethics Committee before the field data collection was conducted. This was done through filling an online questionnaire on what research data was required and data management aspects to be considered during and after the research. This procedure enabled the research to consider ethical aspects when collecting the data, such as ensuring farmers consent to the survey. In addition, how to handle the data after collection, the analysis codes and presentation of the findings with consideration given to any possibilities of the research posing detrimental impact on farmers livelihoods were considered in this research. On data management, a thorough process of collecting and examining the research data was done to ensure its quality, verifiability, and repeatability. This included the creation and submission of a data management plan.

2.4. Software

The following software was used in this research.

- ENVI Classic 5.3 with additional spectral extraction tool developed by ITC Department of Natural Resources for spectral image pre-processing, spectral extraction by polygons (field boundaries).
- MATLAB R2022a for spectra analysis i.e., filtering and plotting graphs.
- ArcGIS 10.4.1for field boundaries data management, maps.
- SPSS for running statistical tests.
- Python (Jupyter Notebook) for building the RF model for feature selection and classification.

3. METHODOLOGY

This chapter describes the steps used in the research to achieve the study objectives. The methodology includes analysis of farmers' interviews and field boundaries data, extraction of spectral signatures from DESIS hyperspectral satellite image, statistical analyses, and classification. The overview of the methodology is shown in Figure 4.

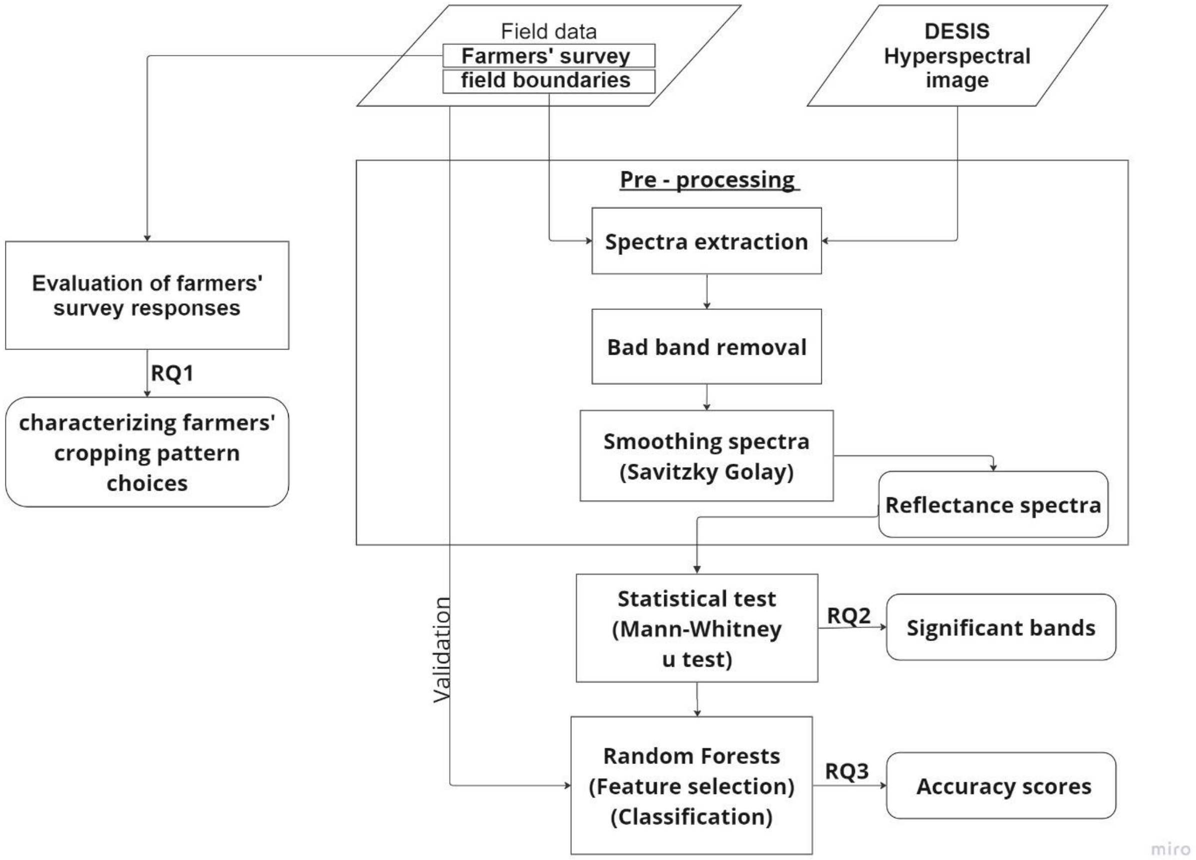

Figure 4: Flowchart of the research methodology.

3.1. Farmer surveys and field boundaries data

The farmer surveys and field boundary data collection were carried out from the 11th of July 2022 to the 5th of August 2022. A stratified random sampling based on the location of the fields to be visited in the four sub-counties was followed. Farmers were interviewed and responses noted by the interviewer (researcher) (see Appendix A for the questionnaire form). To align the research with ethical considerations, the respondents (farmers) were first asked about their consent for the data collection. The purpose of the survey was to derive the crop calendar information of the year of data collection (2022) and the previous year (2021) and to note the factors influencing the choices of the crop types and cropping patterns observed. The influencing factors information was considered during the farmers' survey to better understand farmers' choices and perspectives of cropping practices and to compare the findings with previous literature. In addition, field boundaries were delineated as polygons using a handheld Global Positioning System (GPS) device with a horizontal error of ±3m. The field polygons were acquired to be later linked to the DESIS hyperspectral image for spectral analysis. Every field boundary measurement was linked to the corresponding farmer's interview IDs. Geotagged pictures and dates were taken and indicated in the survey form for further inspections later in the research.

3.1.1. Field data analysis

314 field boundaries were taken during the field data collection from different locations within the study area. Out of these 314 fields boundaries collected, 226 farmers provided survey responses across the four sub counties. From the 226 farmers survey responses, 175 responses pertained the monocropped and intercropped fields responses that were further used for evaluating the factors that influence choices of cropping patterns. For the field boundaries, the polygons were overlayed over the DESIS hyperspectral image using ArcGIS software. Unfortunately, 100 field boundaries were not covered by the satellite image boundaries as shown in Figure 6. This means a total of 214 fields remained within the image coverage. Since the image used was from 2021, the crop calendar information based on farmers' response for 2021 were evaluated. In addition, the size of the fields that were above 900 square meters (0.1ha) were considered for further analysis because of the relation to the size of the pixels of the image used.

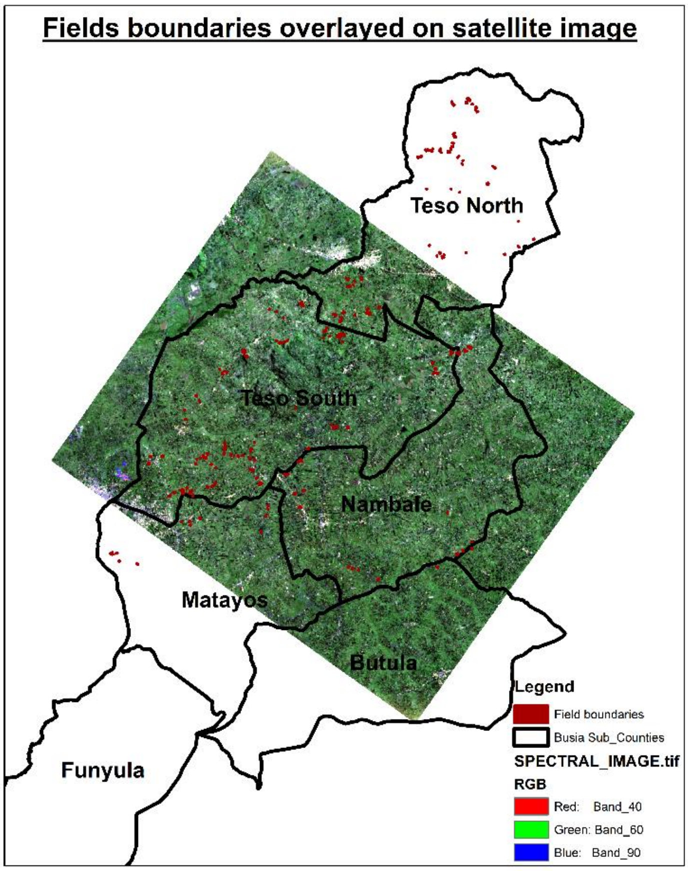

Figure 5: The position and coverage of the satellite imagery and the fields that were excluded for further analysis.

From the 214 polygons previously mentioned, only 74 fields of monocrop and intercropped fields information were retrieved based on the above criteria and could be further used for spectral analysis. Out of the 74 fields, 50 were monocrop maize fields (maize) and 24 were intercrop maize fields (imaize), where beans and/or soyabeans were interplanted. When further evaluating the farmer's responses per county, 6 fields of maize and 4 fields of imaize were not clearly reported by the farmers as either monocrop or intercrop fields. These field plots were then excluded from further analysis because they could not be assumed to belong in either cropping pattern class. Therefore, the final monocropped maize fields considered in relation to the image spectral analysis were 42 fields while intercropped maize fields were 20. In total, 62 fields were then used in the subsequent spectral analysis. Table 2 provides a summary of the number of fields in each sub-county, and Figure 6 shows the location of the fields that spectral signatures were drawn. Few fields had regular shapes and most of them were irregularly shaped.

Table 2: Distribution of fields per subcounty level.

Sub county	**No. of Farmers**	**No. of maize monocrop fields**	**No. of maize intercrop fields**
Teso North	14	11	3
Teso South	32	20	12
Nambale	13	9	4
Matayos	3	2	1
Total (n)		**42**	**20**

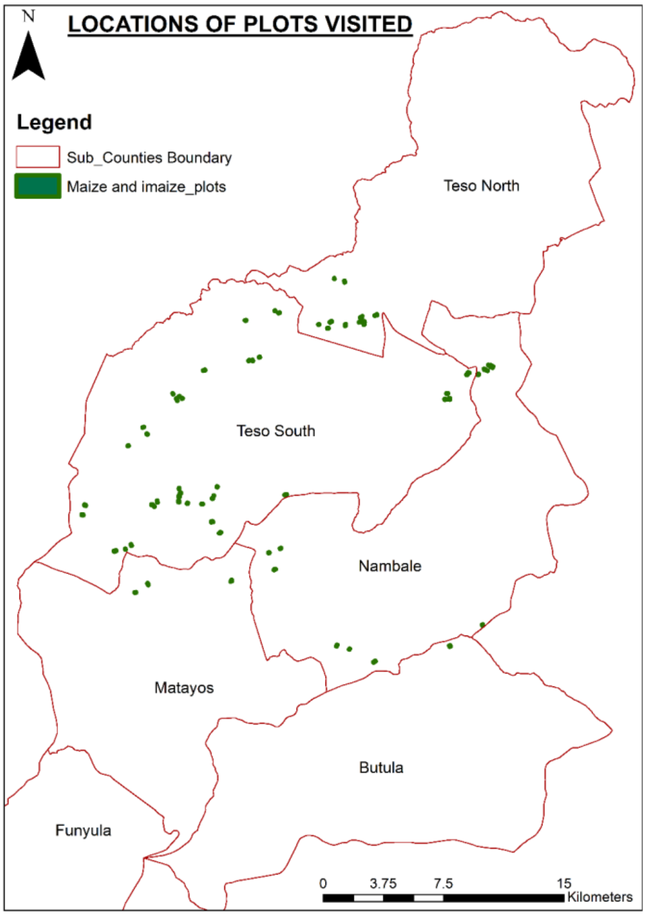

Figure 6: Final maize monocrop and maize intercrop field locations used in this study.

3.2. Image pre-processing

One hyperspectral remote sensing image (DESIS-HSI-L2A-DT0595647612_002-20210607T064016-V0215) was used for the study. Since the DESIS image was already atmospherically corrected, few pre-processing steps were performed as described in the following subsections. ENVI software was used to perform these pre-processing steps, and then MATLAB was used for further spectra analysis.

3.2.1. Polygon spectra reflectance extraction

The spectrum extraction tool was used to extract a spatially aggregated spectral profiles delimited by field boundaries. It uses polygon features, to extract spectral profiles from the spectral image and applies a spatial aggregation function to the profiles (Figure 7). The 62 field polygons as mentioned previously were overlayed on the image, and their respective reflectance spectra were extracted using the spectrum extraction tool in ENVI software. Only pixels completely within the boundaries were extracted and used for spectral analysis. The output of this process was written to an Excel file with each spectrum per field as columns and polygon field IDs chosen as rows.

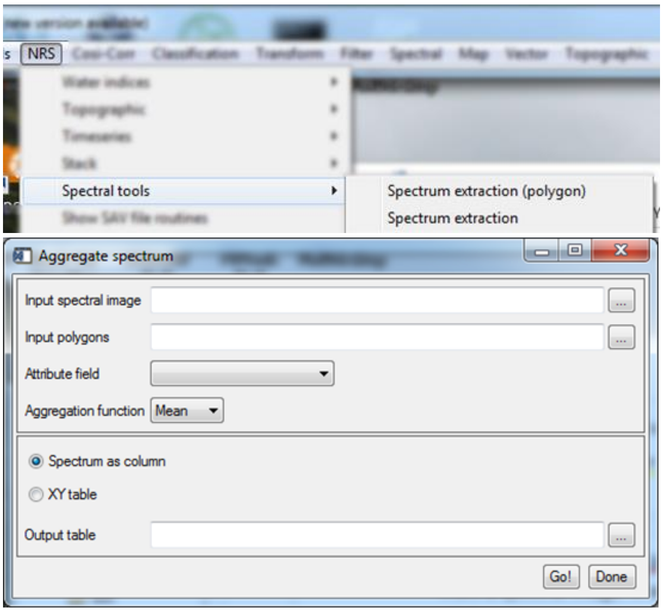

Figure 7: User interface for spectrum extraction using NRS Spectral tools.

3.2.2. Removal of bad bands

When analysing the spectral profile of the fields with all the 235 spectral bands, bands 1 to 7 were removed due to noise with reflectance below zero. 228 useful bands remained for the study with wavelengths ranging from 419.4nm to 999.5nm as shown in Figure 8.

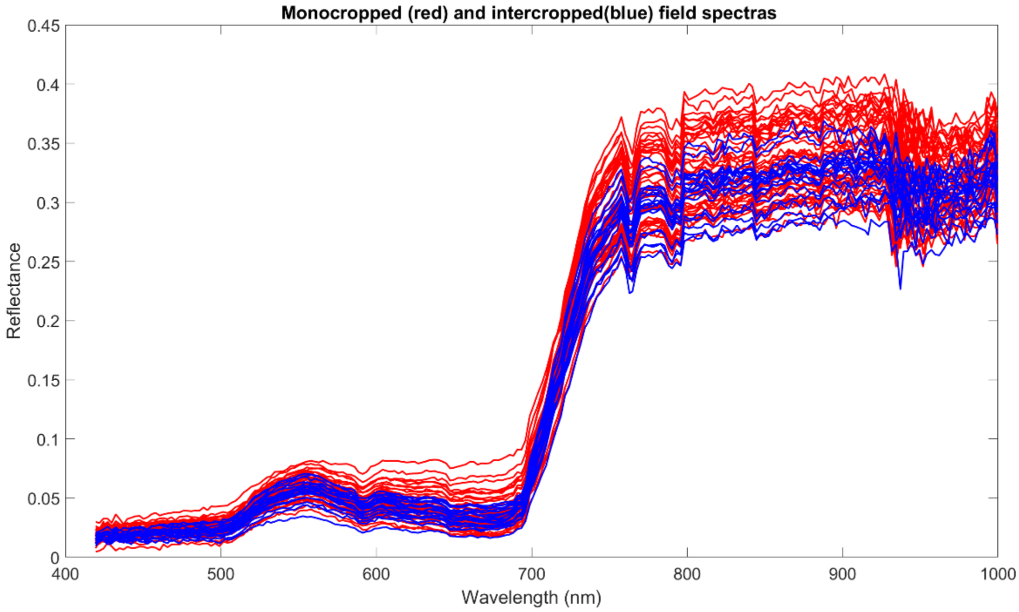

Figure 8: Spectral signatures of maize (red, n =42) and imaize (blue, n = 20) fields after removal of noisy bands (400 – 418nm).

3.2.3. Spectra smoothing: Savitzky-Golay filter

Smoothing of the spectra is a crucial step in hyperspectral remote sensing. Smoothing is required to minimize random noise, which is done by reducing the difference between the individual pixel intensities from the neighbouring pixels (Lowe et al., 2017; Zhao et al., 2020). Savitzky-Golay smoothing was performed using a window frame size of 5 and second degree of polynomial as these parameters were identified ideal for smoothing the spectras as seen in Figure 8. Savitzky-Golay filter was used as it removes the noisy data by fitting polynomials to a subset of data (a spectrum), then evaluates the polynomial at a specific single point to smoothen the signal (Lowe et al., 2017). In addition, Savitsky-Golay filter was considered because it preserves high-frequency signal components without losing information (Jardim & Morgado-Dias, 2020), that is, ideally removing noise without altering the spectral features.

3.3. Statistical tests

To test whether there is a significant difference in the narrowbands of hyperspectral data between two independent samples (in this case, 42 samples of maize monocrop fields and 20 samples of imaize fields), either a two-sample *t*-test or a non-parametric test such as the Mann-Whitney *U* test can be used (Kumar et al., 2019; Manevski et al., 2012). In this case, the purpose of the test was to identify the most significant bands by comparing the distribution of reflectance values in each of the 228 bands between maize and imaize fields (Manevski et al., 2012; Sobhan, 2007). A *t*-test requires that the distribution of reflectance values are normal. While this may be the case for some bands, it is unlikely to be the case for all bands. On the other hand, the Mann-Whitney *U* test does not assume any particular distribution for the data (Kumar et al., 2019; Schmidt & Skidmore, 2003). In this study, given the small and unequal sample sizes, and the number of bands (228) to be tested, the Mann-Whitney *U* test was chosen to determine the significant differences for each of the 228 bands with a *p*-value <0.05 as the cut off for determining significant differences (Shafri et al., 2011). SPSS software was used for this part of the methodology.

3.4. Feature selection and classification using RF

RF is an ensemble classifier that is suitable for the classification of hyperspectral data as it is known to handle the high dimensionality of data (Belgiu & Drăgu, 2016; Prospere et al., 2014). It uses randomly selected subset of training samples and does not consider any underlying probability distribution for input data (Kale et al., 2017). The advantage of using RF classification is that it produces good classification results in terms of accuracy values, and it is insensitive when it comes to training labels (Kale et al., 2017; Liaw & Wiener, 2002). RF requires optimization of user-based parameters in the model to guarantee the high accuracy of classification as well as reproducibility of the results when different or several runs are performed on the model (Richard et al., 2017; Vaiphasa et al., 2005). In this study, the optimal number of trees (^{n}tree), maximum variables used at each tree (^{m}try), and random_state of RF were optimized. For the ^{n}tree and ^{m}try, this was done by inputting the data to the model, split into training and testing set and finding the best combination of ^{n}tree and ^{m}try that gives the highest accuracy on the testing set. The final RF model used ^{n}tree of 100 and ^{m}try of 10. A random_state of 0 was used as this parameter is crucial for reproducibility. Python was used for this section of the methodology.

3.4.1. Feature selection

One common approach to selecting a subset of bands for discrimination is to use feature selection or feature reduction methods. Feature reduction analysis is crucial to obtain uncorrelated wavelengths from the significant wavelengths selected after the statistical test (Kumar et al., 2019). After statistically significant wavelengths were identified, feature selection was performed to reduce the number of significant bands by identifying a subset of relevant features that can improve the performance of a classification model, while discarding redundant that can lead to overfitting or decrease the model's interpretability (Belgiu & Drăgu, 2016). In this study, the RF out of bag method (OOB) was adopted. This method was used to calculate the importance of specific predictor variables (wavelengths) that will aid in discriminating the cropping patterns (Adam et al., 2012; Prospere et al., 2014). The first step was to use about two thirds of the samples in the training set for training and the remaining samples for error assessment. In the next step, the importance of each variable (each significant wavelength) was randomly permuted and was passed down in every tree to get new predictions while analysing the mean error for each decision tree. In the end, mean difference before and after the premutation was calculated, and a ranking index was used to identify the variables(spectral bands) with the largest importance in the classification process (Liaw & Wiener, 2002; Prospere et al., 2014).

Further, a simple step forward feature selection technique was used to find the best optimal subset of wavelengths for classification (Adam et al., 2012). Step forward feature selection uses a sequential feature selection that starts by evaluating each feature and selects the best performing selected variable determined by the evaluation criteria, e.g., accuracy or misclassification error. The step continues for all possible combinations and subsequent features are evaluated and a feature added and so on until the required number of features are selected. In this model, 5-fold cross validation was performed to select the best features (spectral bands) for classification.

3.4.2. Classification and model validation

The 62 samples were used as training and testing datasets for the RF classifier. 70% of the data was used to train the model while 30% was used as a test set as shown in Table 3, representing each class pattern. A confusion matrix was constructed to assess the accuracy of the classification performance. From the confusion matrix, the overall accuracy (OA), producer accuracy (PA), user accuracy (UA), and Kappa index (score) were calculated for the algorithm used to classify the patterns. The F1 score was calculated to assess the class accuracy and give equal importance to precision and recall by combining the PA and UA into a fused measure (Richard et al., 2017).

Table 3: Number of samples used for training and testing sets

Class	Number of samples	Training set 70%	Testing set 30%
maize	42	29	13
imaize	20	14	6

4. RESULTS

This section covers the findings from the study including the data obtained during the field visit, spectral signatures analysis of the cropping patterns, results of statistical tests, and classification results.

4.1. Field data: Descriptive analysis of farmers' survey

Farmers from four sub-counties in Busia (Figure 6): Teso North, Teso South, Nambale and Matayos were visited and interviewed on crop calendar information as well as factors affecting their choice of cropping pattern. Regarding the crop calendar information for 2021 and 2022, most farmers planted in February/March and harvested in June/July. These planting and harvesting months in the region for the two cropping systems align with the seasonal rainfall data shown in Figure 2 and the county information about the region cropping calendar (Busia, 2018). Based on field data collected, the planting and harvesting times are summarized in Tables 4 and 5 and Figures 9 and 10 for monocropped and intercropped fields. Based on the combined survey responses of 2021 and 2022, the total number of farmers' responses on factors influencing choice of cropping patterns, that is, monocropped fields were 88 while intercropped farmers' responses were 87, summing up to 175 farmers' responses.

Table 4: Number of monocrop farmers and the planting and harvesting months.

	Planting				Harvesting			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Maize monocrop	3	22	35	28	2	20	40	26

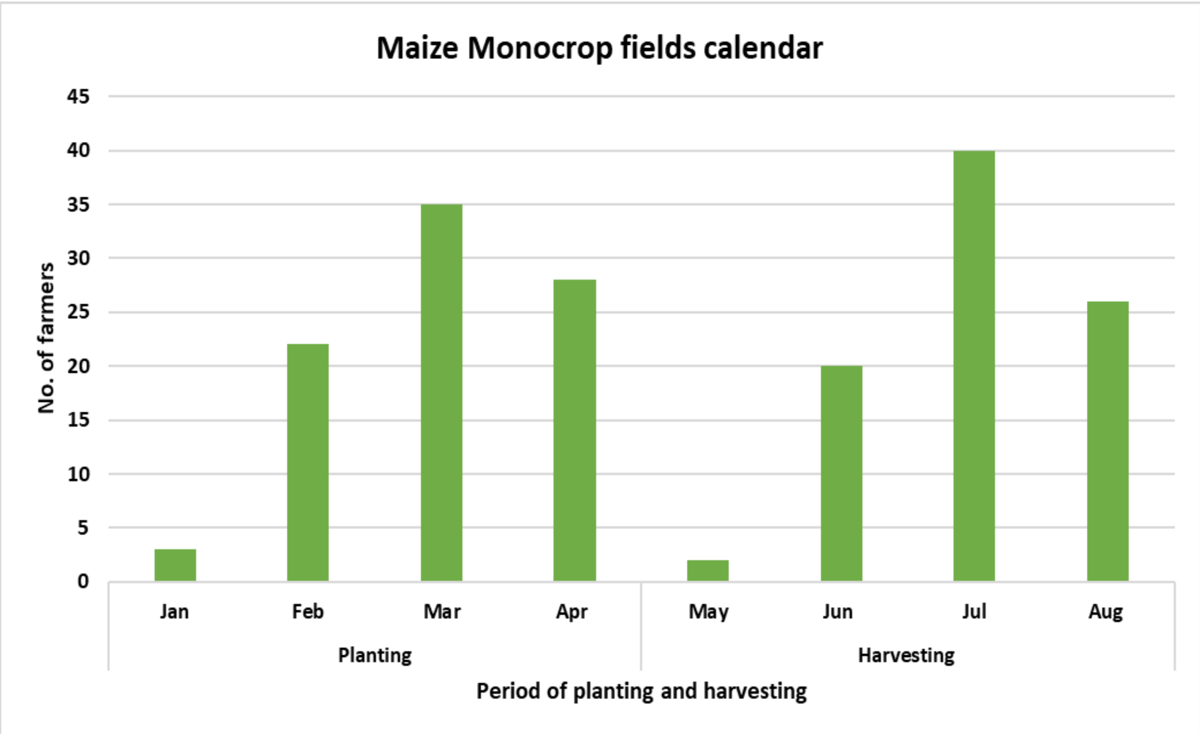

Figure 9: Planting and harvesting months of monocropped fields.

Table 5: Number of intercrop farmers and the planting and harvesting months.

	Planting				Harvesting			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Maize	2	13	43	29	2	14	45	26
Beans/soyabeans	-	15	42	30	14	43	23	7

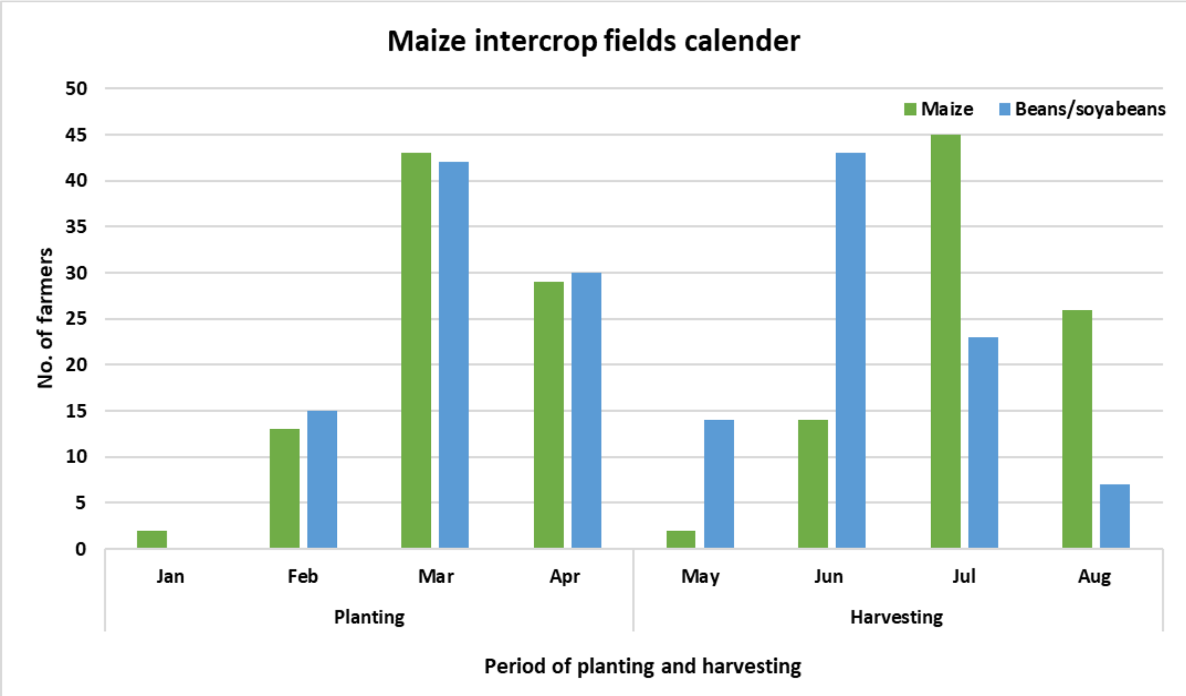

Figure 10: Planting and harvesting months of intercropped fields.

Further, from the farmers' survey during the field data collection, the farmers were asked about the factors influencing their cropping pattern choice. Table 6 summarizes the factors that most farmers highlighted and shows the number of farmers' responses to each factor. More explanation is provided in the discussion section in Chapter 5.

Table 6: Factors influencing the choice of cropping pattern based on farmers' survey data.

	Factors	No. of farmers Intercropping	No. of farmers Monocropping
1	Size of the farm (field)	21	-
2	Personal experience (practice over the years)	14	37
3	Subsistence use (Family consumption)	16	16
4	Availability of resources (machinery/farm inputs/labour)	12	9
5	Market demand (Financial support/boost)	11	26

6	Pest prevention /symbiosis	13	-

4.2. Spectra smoothing

Figures 11 and 12 show the resulting smoothened signatures after Savitsky-Golay filter was used and the average spectra of 42 maize fields and 20 imaize fields, respectively.

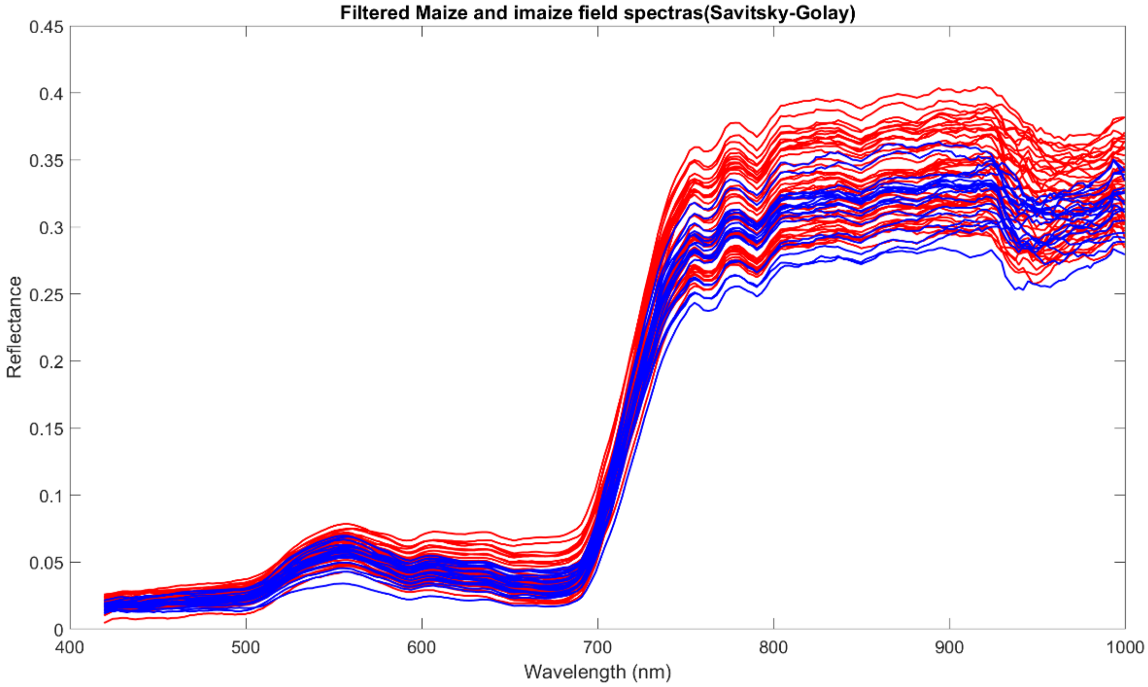

Figure 11: Spectral signatures of maize (red, n=42) and imaize (blue, n=20) fields after smoothing using Savitsky-Golay filter with a window size of 5.

After filtering all the fields, the average of 42 reflectance spectra from maize and 20 reflectance spectra from imaize fields was calculated ,and Figure 12 shows the result of the averaged spectra.

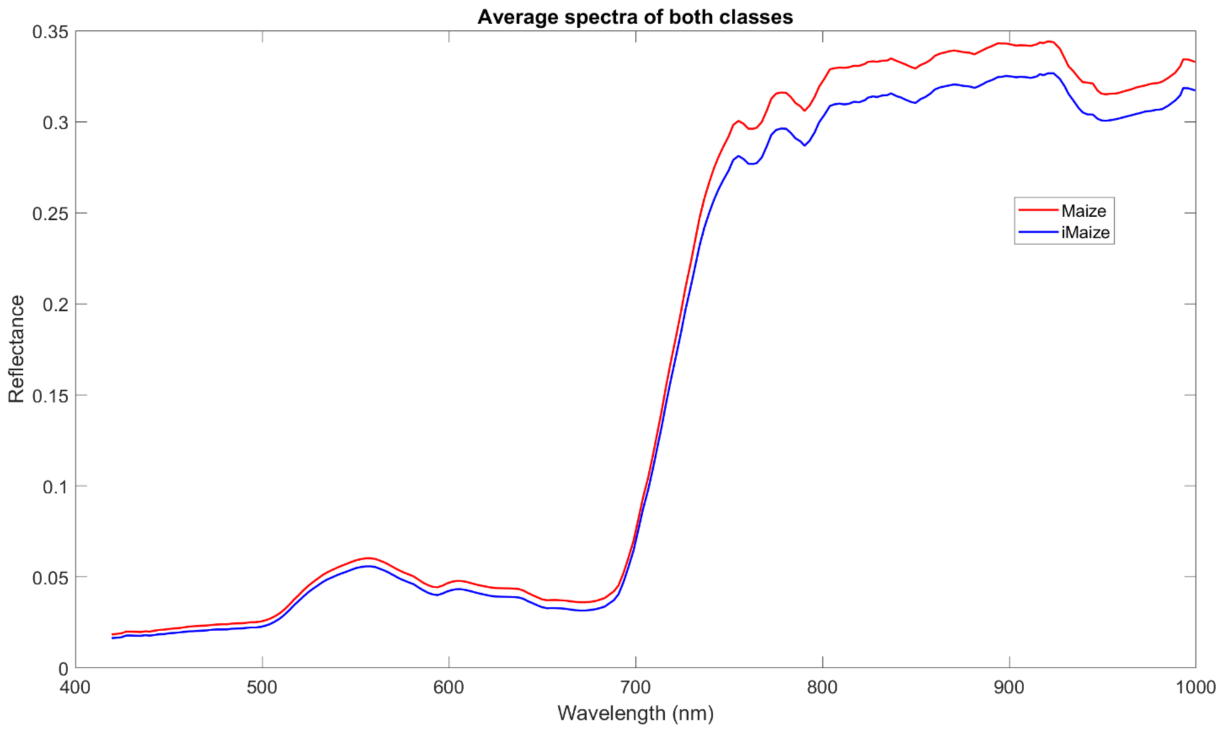

Figure 12: Average reflectance spectra for maize and imaize fields (maize = 42 fields, imaize = 20 fields).

It can be observed from Figure 12 that monocropped fields (maize) have a higher reflectance than intercropped fields (maize) and that from the red edge towards the NIR region of the electromagnetic spectrum, the maize and imaize reflectance shows a relatively wide spectral difference as compared to the visible part of the spectrum.

4.3. Results of Mann-Whitney *U* test

To check the distribution of the data at different locations in the wavelength, a box plot was visualized as shown in Figure 13. The box plot showed a relative similar spectral curvature to the average spectra shown in Figure 12 above.

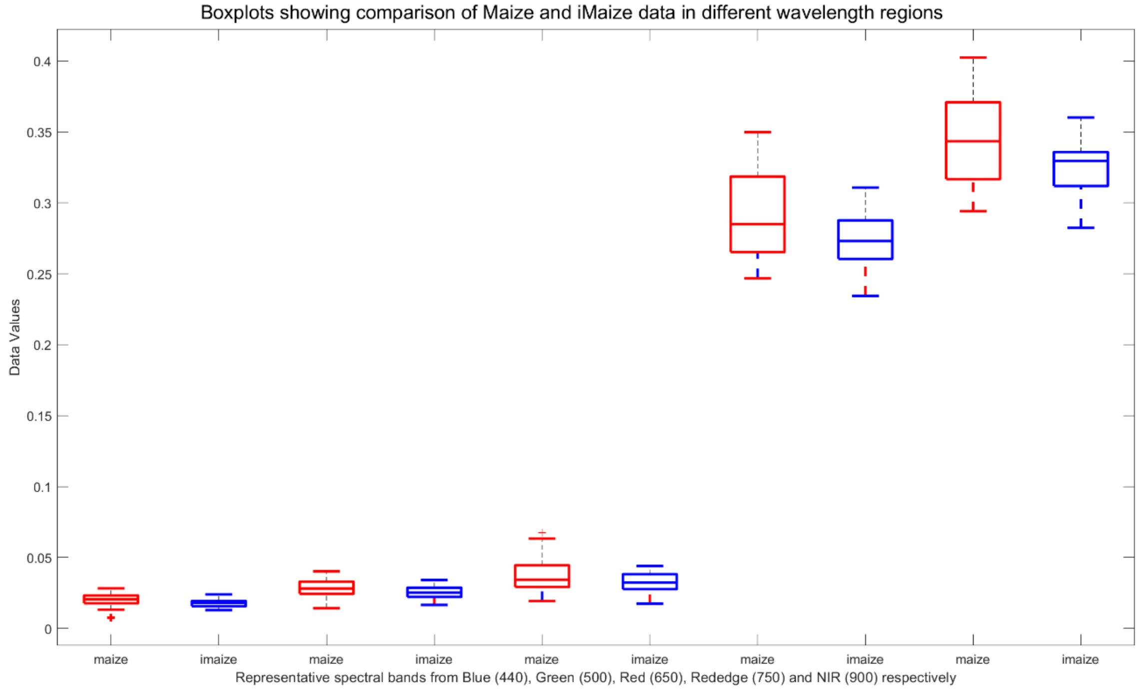

Figure 13: Box plot showing the variations of maize and imaize reflectance data in different wavelength regions

Using the 42 samples of maize monocropped and 20 samples of maize intercropped fields with the Mann-Whitney *U* test, the comparison of the median reflectance between the two groups showed a statistical difference from one another. One hundred and ten (110) optimal narrow bands from the visible region to the NIR region of the spectrum showed significant differences (with p-values < 0.05) as depicted in Table 7. Results of individual bands statistical p-values are provided in Appendix B. The interpretation of the results is further discussed in Chapter 5.

Table 7: Results of the Mann-Whitney U test showing the number of significant bands in each wavelength region.

Wavelengths region (nm)	Region	No. of Significant bands
419 - 600	Visible	35
700 - 800	Red Edge	25
801 - 1000	Near Infrared	50

4.4. RF classification and accuracy assessment

When the RF was adopted with the optimized parameters together with the feature selection method, 5 optimal bands were selected from the 110 bands from the statistical results, to be used for further classification purpose. The selected bands from the feature selection were dominated in the red edge and part of NIR region (752.2nm, 767.5nm, 775.2nm, 783nm, 814.2nm). This can be reflected from the average spectra graph (Figure 12) and the box plot (Figure 13) that shows a wide difference of the maize and imaize spectra. Using the 5 bands, the RF classifier was able to classify the cropping patterns with 74% overall accuracy as shown in Table 8. Evaluating the producer and user accuracy of the two cropping patterns, maize class attained a producer accuracy of 71% and 80% for the imaize class. The user accuracy was 91% for maize and 50% for the imaize, indicating a 50% chance of imaize class being misclassified as maize. The F1 score for maize class was 80% and 62% for imaize class. The kappa coefficient of agreement for the two cropping patterns was low at 0.43 indicating the complexity of classifying these cropping patterns. The classification results are summarized in Table 8 and also given in Appendix C from the code result. The interpretation of these results is further discussed in Chapter 5.

Table 8: Classification results of maize and imaize fields using RF classifier.

Class	Producer accuracy	User accuracy	F1 score	Overall accuracy	Kappa score
maize (monocropped fields)	71%	91%	80%	**74%**	0.43
imaize (intercropped fields)	80%	50%	62%		

5. DISCUSSION

This study evaluated farmers choices on maize cropping patterns practices, i.e., maize monocropping and intercropping patterns as a common practice for most small-scale holder farmers in Busia County, Kenya. A total of 175 farmers' responses from field data collection were evaluated to determine the common factors that farmers perceive to influence their choices for cropping patterns. It further examined the potential of using DESIS hyperspectral satellite data to discriminate between these cropping patterns. Based on the image area coverage, responses on farmers survey in 2021 and size of the fields, 42 monocropped fields and 20 intercropped fields boundaries were further used to extract reflectance spectra from DESIS hyperspectral image for analysis. Section 5.1 and 5.2 provide a detailed discussion of the results. Section 5.3 provides recommendation based on the findings of the research.

5.1. Evaluation of factors influencing farmers' choice of cropping pattern

The reasoning underlying farmers' responses on factors influencing the choices of cropping pattern is discussed in this section with comparison and contrast to other literature. Relative to the hypothesis, this section supports the alternative hypothesis highlighting the physical and social-economic factors being the most influencing factors on farmers' decisions on cropping patterns and land management practises as observed from other literature as well.

Size of the farm (fields)

Table 6 reveals that the size of the farm is one of the major factors that influence farmers to practise a particular cropping pattern. According to the farmers exercising intercropping, utilizing the small piece of land is beneficial to them because they get to harvest more than one crop from the piece of land. Thus, according to them, the small piece of land is utilized fully to give them maximum production for subsistence and partly commercial use. A study in China by Hong et al. (2020) highlights that farm size has an impact on the choice of cropping patterns. The study indicates that the size of a farm has a significant impact on yield and profit gains for the farmers. For example, farmers practising intercropping tend to have higher yields and profit per unit of land than farmers practising monocropping. In a study by Mogaka et al. (2021), farmers embrace crop diversification in intercropping to manage climatic risks. Interestingly, in several studies (Beyene et al., 2019; Hong et al., 2020), this factor has been closely related to household needs i.e., family requirements, as is the case with this study. Farmers indicated that practising intercropping helps to improve household food security, their explanation being that, as they wait for the main crop to mature, the other crop can sustain the family.

Availability of resources

Availability of resources, in this case, includes the financial capability when it comes to purchasing farm inputs and hiring machinery and labour that would increase farm efficiency. In terms of fertilizers for example, most of the farmers in Busia complained about the prices of fertilizers being way above their means hindering them to apply to their crops and that led to low production levels. When it comes to machinery, few of the farmers indicated to have used machinery for example, tractors to plough their land. This is true mostly for farmers with relatively large farms. Few used oxen to prepare their farms while most of them used their own manual labour with forks. Most farmers who practise intercropping indicated to use manual labour more because of the cropping pattern dynamics. A study by Hong et al. (2020) indicates that agricultural machinery is an important element that affects the choice of monocropping and intercropping. This compares to a study in Tanzania by Greig (2009) who discusses the responses of farmers on the availability of machinery and fertilizers, stating that these availabilities of resources are a fundamental factor that influences farmers' choices. When it comes to labour, it is expected to influence both cropping practices positively. With intercropping, more labour is required as more work in planting and managing the crops is required compared to monocropped farms. This compares to the study by Hong et al. (2020) in China, who indicates that household size also significantly contributes to the labour requirements.

Personal experience (practice over the years)
Another key finding was the farmer's experience in the years they have practiced a particular cropping pattern. Based on years of experience in farming, most farmers practice what they term as "norm" or "tradition." One subsistence farmer stated, "it is our tradition to do intercropping, even our parents did it." In this case, the preference was evidently influenced by what has been practised over the years and handed over from generation to generation. This is compared to a study by Briggs (1985) in Sudan who indicated this as one of the top three decision factors as it shows a satisfactory sense of knowledge of what has been done before. However, from the field visit, it was noted that most farmers still have poor management practises in their lands and therefore, an emphasis on the significance of agricultural extension services is required (Briggs, 1985; Mogaka et al., 2021). Extension services educate farmers on best practices, especially new methods and techniques that can help them maximize their farm output.

Market demand
Farmers practising both cropping patterns i.e., mono- and intercropping stated that apart from growing the crops for subsistence use, they considered financial advantages when they sell their produce. Most of the farmers responded in depth that market demand for their surplus products helped them offset some financial burdens that pressed them, including the loans taken to purchase farm inputs and other financial needs such as paying school fees for their children. A study by Briggs (1985) reveals that farmers indicated that the reliance of market for the crops produced, as well as the income generated from selling the crops, helped farmers' financial needs. This factor shows that practising both cropping patterns have potential benefit to the farmers in terms of attaining financial security and improving their living standards when they have surplus farm production that meets the market demand.

Pest prevention/plants symbiosis
The other important factor from farmers' responses is that intercropping reduces the risk of total crop failure when there is an outbreak of pests and diseases compared to monocropping (Hong et al., 2020; Matusso et al., 2012; Mthembu et al., 2019; Robbins, 2022). Although literature and research indicate that growing more than one crop in the fields helps in pest infestation, not all farmers shared the same perspective in this research's survey. From the survey, most of the farmers who practice intercropping indicated that they mainly plant beans with maize because beans act as manure to the maize once it is harvested. This agrees with previous studies on intercropping (Hong et al., 2020; Mthembu et al., 2019; Rusinamhodzi et al., 2012; Stomph et al., 2020). According to Stomph et al. (2020), there is an indication that when growing two or more crops, there is a reduction of farm input, however there was a contrast with the farmer's response. Most intercropping farmers indicated that they use more fertilizer when they plant more than one crop, which is why some farmers stick to planting only maize (monocropping) in their farms. In addition, from the field data collection, we noted that several farmers used different types of seeds for maize crops, claiming that some seeds have fast maturity levels and respond differently to weather extremes and to pests and diseases. This is also highlighted by a study in China by Hong et al. (2020), who indicated that farmers stated that different crop species respond differently to weather extremes and diseases and pest outbreaks. However, this aspect requires further research as it was not examined in depth during the data collection and in the literature.

From the evaluation of the factors influencing farmers' choice of cropping patterns, the findings point out to the physical and socio-economic factors as supported by the alternative hypothesis of the research. The findings indicate that the small size of the farms (fields), large household dependency, farmers' experience and pest prevention measures are the major factors that influence farmers to practice intercropping patterns. When it comes to monocropping practices, farmers considered mostly household needs, availability of resources such as machinery and farm inputs and the availability of market demand to sell their surplus farm products. Most of these physical and socio-economic factors compare to what previous literatures have discussed when evaluating farmers' choices of crop type establishment or cropping patterns in different regions that focus especially on small-holder farmers. However, some of the farmers' narrative differed with what was in the literature, for example, the use of fertilizers where farmers who practise intercropping strongly narrated that more fertilizer is required when planting two crops in the fields while literature indicates that intercropping practises enabling crops to benefit from fertilizer input of one crop. It is therefore important to take these insights into consideration for future studies and when designing interventions for food security.

5.2. Assessing the potential of DESIS hyperspectral imagery for cropping pattern discrimination.

Maize cropping patterns were characterized by examining whether there were significant differences in the spectral signatures of the fields using individual bands comparison using a Mann-Whitney *U* test and then bands with optimal characteristics were further used for evaluating the classification. The results from average spectra (Figure 12) and boxplot visualization (Figure 13) indicated that monocropped (maize) fields have a higher reflectance than intercropped (imaize) fields, and that a major difference in the spectral signatures could be observed from the red edge and NIR regions of the electromagnetic spectrum. When the Mann-Whitney *U* test was performed, the results from Table 7 related with was observed in Figures 12 and 13, where a smaller number of bands were from visible (blue and green) region compared to the red edge and the NIR region with a greater number of bands. This indicates that the maize monocropped fields can be statistically distinguished from maize intercropped fields majorly using the bands from red edge and the NIR regions of DESIS hyperspectral satellite data. From statistical analysis, the results of the Mann-Whitney *U* statistical test hereby reject the null hypothesis and support the alternative hypothesis there is a statistical difference in the spectral bands and that the red edge and NIR region give the optimal bands for crop pattern discrimination, indicating the potential discrimination of the maize based cropping patterns. The results agree with previous studies (Darvishzadeh, 2008; Prospere et al., 2014; Schmidt & Skidmore, 2003; Thenkabail et al., 2004; Vaiphasa et al., 2005) that have demonstrated the red edge and NIR bands being suitable for species and plant communities discrimination.

When it comes to classification, the results indicated moderately good classification accuracies, where overall accuracy of 74% was attained despite the low number of samples used for testing. Monocropped fields (maize class) showed a producer accuracy of 71% and a user accuracy of 91%. Intercropped fields (imaize class) gave a producer accuracy of 80% and user accuracy of 50%. The user accuracy of imaize indicated the complexity of performing classification for intercropping as the reflectance is more likely to be identified as maize fields. Further, the Kappa coefficient was relatively low (0.43) compared to most studies that reported Kappa values close to 1 (Adam et al., 2012; Richard et al., 2017; Vaiphasa et al., 2005). The low Kappa coefficient from the results indicates the complexity of classifying maize based cropping patterns. Several limitations in the research could aid in the interpretation of the classification results obtained, particularly the kappa coefficient and the user accuracy of imaize class. They include;

i. The farm(field) management practices of the farmers and the complexity of the nature of the cropping patterns i.e., monocropped (maize) and intercropped (imaize) fields, as observed during the field data collection. It was noted that some farmers had poor land management practices and that some individual fields, especially monocropped maize fields, contained a lot of weeds. In this case, the spectra reflectance of such fields could give similar characteristics to an intercropped field and moreover, intercropped field spectral signature could be easily misclassified for a monocropped field because of similar spectral reflectance. Farm level challenges, therefore, could have a possible influence on the discrimination and classification power of the spectral bands that can lead to misclassification and low accuracy levels.

ii. The number of samples. During the field data collection, a number of challenges were encountered during the implementation. To begin with, to engage the local farmers in the regions, the team was required to inform the all the local authorities in the sub-counties level and hence get clearance to collect the data. This took several days to achieve, reducing the number of data collection days. The major challenge was the sizes of the fields in the study area (most fields were below 0.1ha) which could not be used in the research. This factor reduced the number of data samples especially for intercropped fields. Moreover, given the literacy level and the age of most farmers, crop calendar information for the 2021 deemed a challenge in terms of memory, thereby resulting in less information, which resulted to less sample size, especially for imaize class used for spectral analysis. In addition, the DESIS image used did not cover all the field data points collected hence reducing the number of samples used in the study.

iii. The satellite image used in this study. There were quite some challenges with the image used for analysis. The time of acquisition of the images (DESIS hyperspectral acquired in June 2021 and PRISMA image acquired in July 2022) were towards the harvesting stage of the crops based on the study area crop calendar, which was not optimal. From the literature, the optimal acquisition dates of imagery are during stem elongation and flowering crop development stage. A study by Arvor et al., (2011) and Richard et al., (2017) points out that, at stem elongation and flowering development crop stage, monocrop and intercropping field patterns can be more distinguishable than at later growing stages. After flowering for example, the two maize cropping patterns seem to have same morphological and spectral properties. Moreover, the spatial resolution of the DESIS hyperspectral image, which is 30m resolution was used, whereas a very high spatial resolution of less than 5m would be preferable (Richard et al., 2017) for such complex cropping pattern analysis. Previous studies by Rebecca (2020) and Richard et al. (2017) explain in depth the complexity of discriminating and classifying cropping patterns using multispectral data.

From the hypothesis, these classification results were relatively lower than expected, considering the utilization of red edge and NIR regions of the electromagnetic spectrum. Previous studies have utilized the same regions and the classification results are high (OA > 80% with Kappa values close to 1). However, the limitations stated above could have negatively influenced the expected outputs. Despite these challenges, hyperspectral remote sensing shows potential for discriminating and mapping cropping patterns as a contribution to food security-related research questions and thus requires further exploration with appropriate image acquisition dates, appropriate spatial and temporal resolution, and sufficient ground data.

5.3. Recommendations

Based on farmers responses to the factors stated above, it is apparent that physical and more so, socio-economic factors influence farmers' decisions on the choice of cropping patterns. It is, therefore important to note that understanding these factors and the decision dynamics of farmers is crucial in designing site-specific sustainable farming interventions and policies. This study contributes to that understanding. From the use of remote sensing data, the result of this study show that DESIS hyperspectral remote sensing can be used to identify different maize cropping patterns for mapping and monitoring of intercropping patterns, but with recommendations on the appropriate timing of the imagery to coincide with field conditions where the main crop and intercropped crop are both visible.

This work can further be used as a decision support framework. For example, when the cropping patterns are accurately classified and then mapped, the information can be used to inform agriculture stakeholders and even the governments where cropping patterns occur and can thus enable extension support to farmers on how to maximize their production through the provision of agronomic support, including fertilizers. In terms of food security assessment, the potential contribution of intercropping to food production and its contribution to climate change and ecosystem impacts can be assessed. The results cannot only be used at a local scale but can be upscaled to a regional scale, national and global scale to inform on these agroecological practices. Moreover, land use policy makers can also use the information to develop a system of incentives to farmers adopting sustainable agriculture as more food is produced while maintaining the ecosystem.

6. CONCLUSION

This study aimed to evaluate the choices that influence farmers to practice a particular maize based cropping pattern and assess the potential of using DESIS hyperspectral remote sensing to discriminate and classify these cropping patterns. To do this, three objectives were set and achieved using field data collection which included farmers' survey and field boundary measurements with DESIS hyperspectral satellite data. The first objective was to evaluate the factors that influence farmers to practice different cropping patterns. From the survey responses evaluation, small size of the farms (fields), large household dependency, farmers' experience and pest prevention measures are the major factors that influence farmers to practice intercropping patterns. When it comes to monocropping practices, farmers considered mostly the availability of market demand to sell their farm products, household needs and availability of resources such as machinery. This indicate that socio-economic factors play a significant role in influencing farmers choices and hence it is important to understand these factors for targeted policy interventions.

The second and third objectives were to assess the potential of using DESIS hyperspectral satellite data to discriminate the two maize based cropping patterns and analyse the classification accuracy, respectively. The discrimination was based on spectral analysis involving identification of spectral bands that are statistically different from the spectral signature of monocropped and intercropped maize fields. Using a non-parametric statistical method, the Mann-Whitney *U* test, 110 spectral bands from the visible, red edge and NIR region showed significant difference. The result thus supported the alternative hypothesis that monocrop and intercrop patterns are different. For the classification, the study aimed at identifying the best optimal bands that can be used for classification. Using RF feature selection, 5 informative bands from the red edge and NIR region were identified (752.2nm, 767.5nm, 775.2nm, 783nm, 814.2nm) and used for classification. An overall accuracy of 74% was achieved, with user accuracy of 91% for maize and 50% for imaize classes, showing potential land cover classification of the maize based cropping patterns using DESIS hyperspectral remote sensing but recognising the limitations due to the DESIS image coinciding with the later growing stages of the maize crop which limited classification accuracy for imaize.

This research is the first attempt to understand the factors influencing farmers choices of cropping pattern and also use to use DESIS hyperspectral remote sensing to discriminate the maize cropping patterns in the complex and heterogenous landscape in Busia, Kenya, and sub-Saharan Africa in general. Hence, there were no directly comparable prior studies to put the classification accuracies into context. Despite this gap and the limitations discussed, the analysis of the study and the results show potential in discrimination of cropping patterns using hyperspectral satellite data. However, it is important to note that this process is complex and challenging especially in heterogenous landscapes, and that field surveys that aid in understanding farmers land management practises can play a critical role in the interpretation of results. Thus, the exploratory nature of this research has opened more avenues for future research.

In conclusion, it is important to understand and involve stakeholders' perspective, in this case the farmers, by understanding the factors that influence their decisions to different agricultural and land use practices. The results of this study further highlight that discrimination of maize monocropped and maize intercropped fields can be challenging but with the use of early season hyperspectral data, this can be made possible. Thus, in the context of food security, understanding the socio-economic factors that influence farmers choices can aid in providing necessary support that is required to maximize production of food to feed the growing population both locally and globally. In addition, the new and emergent spatial technologies, and particularly hyperspectral remote sensing, plays a significant role in informing the future of our food production and are expected to improve the monitoring of agricultural management practices such as cropping practices (e.g., irrigation) and cropping patterns (e.g., monocropping and intercropping) in the challenging context of food security.

LIST OF REFERENCES

Adam, E. M., Mutanga, O., Rugege, D., & Ismail, R. (2012). Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and Hyperspectral data resampled to HYMAP. *International Journal of Remote Sensing*, *33*(2), 552–569. https://doi.org/10.1080/01431161.2010.543182

Adam, E., & Mutanga, O. (2009). Spectral discrimination of papyrus vegetation (Cyperus papyrus L.) in swamp wetlands using field spectrometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(6), 612–620. https://doi.org/10.1016/j.isprsjprs.2009.04.004

Alonso, K., Bachmann, M., Burch, K., Carmona, E., Cerra, D., de los Reyes, R., Dietrich, D., Heiden, U., Hölderlin, A., Ickes, J., Knodt, U., Krutz, D., Lester, H., Müller, R., Pagnutti, M., Reinartz, P., Richter, R., Ryan, R., Sebastian, I., & Tegler, M. (2019). Data products, quality and validation of the DLR earth sensing imaging spectrometer (DESIS). *Sensors (Switzerland)*, *19*(20), 1–44. https://doi.org/10.3390/s19204471

Arvor, D., Jonathan, M., Meirelles, M. S. P., Dubreuil, V., & Durieux, L. (2011). Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing*, *32*(22), 7847–7871. https://doi.org/10.1080/01431161.2010.531783

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Beyene, A. D., Mekonnen, A., Randall, B., & Deribe, R. (2019). Household Level Determinants of Agroforestry Practices Adoption in Rural Ethiopia. *Forests Trees and Livelihoods*, *28*(3), 194–213. https://doi.org/10.1080/14728028.2019.1620137

Briggs, J. (1985). An exploratory study of farmers' choice of crops in central Sudan. *Transactions - Institute of British Geographers*, *10*(2), 170–180. https://doi.org/10.2307/621821

Busia, C. G. of. (2018). *BUSIA COUNTY INTEGRATED DEVELOPMENT PLAN (2018 - 2021).* pgs 263.

DLR. (2019). *DESIS - Hyperspectral Images - Global.* https://geoservice.dlr.de/data-assets/hxom21uqeo90.html

Eckardt, A., Horack, J., Lehmann, F., Krutz, D., Drescher, J., Whorton, M., Soutullo, M., Engineering, T. B., Drive, S., Box, P. O., German, D. L. R., & Office, D. L. R. W. (2015). *DESIS (DLR EARTH SENSING IMAGING SPECTROMETER FOR THE ISS-MUSES PLATFORM) DLR German Aerospace Center , Institute of Optical Sensor Systems.* 1457–1459.

Greig, L. (2009). An analysis of the key factors influencing farmer's choice of crop, Kibamba Ward, Tanzania. *Journal of Agricultural Economics*, *60*(3), 699–715. https://doi.org/10.1111/j.1477-9552.2009.00215.x

Hong, Y., Heerink, N., & van der Werf, W. (2020). Farm size and smallholders' use of intercropping in Northwest China. *Land Use Policy*, *99*(August), 105004. https://doi.org/10.1016/j.landusepol.2020.105004

Jardim, R., & Morgado-Dias, F. (2020). Savitzky–Golay filtering as image noise reduction with sharp color reset. *Microprocessors and Microsystems*, *74*, 103006. https://doi.org/10.1016/j.micpro.2020.103006

Kale, K. V., Solankar, M. M., Nalawade, D. B., Dhumal, R. K., & Gite, H. R. (2017). A Research Review on Hyperspectral Data Processing and Analysis Algorithms. *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, *87*(4), 541–555. https://doi.org/10.1007/s40010-017-0433-y

Kumar, T., Dutta, D., Chatterjee, D., Chandrasekar, K., Rao, G. S., & Raj, U. (2019). Different Statistical Methods for the Discrimination of Tropical Mangrove Species Using in-Situ Hyperspectral Data. *Journal of Hyperspectral Remote Sensing*, *9*(2), 49. https://doi.org/10.29150/jhrs.v9.2.p49-67

Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest*. *2*(December), 18–22.

Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*(1), 1–12. https://doi.org/10.1186/s13007-017-0233-z

Manevski, K., Manakos, I., Petropoulos, G. P., & Kalaitzidis, C. (2012). *Spectral Discrimination of Mediterranean Maquis and Phrygana Vegetation : Results From a Case Study in Greece*. *November 2020*. https://doi.org/10.1109/JSTARS.2012.2190044

Matusso, J. M. M., Mugwe, J. N., & Mucheru-Muna, M. (2012). Potential role of cereal-legume intercropping systems in integrated soil fertility management in smallholder farming systems of sub-Saharan Africa Résumé. *Third RUFORUM Biennial Meeting 24 - 28 September 2012*, *May 2016*, 1815–1843. http://www.ku.ac.ke/schools/environmental/images/stories/research/potential_role_of_cereal_legume_intercropping_systems.pdf

Mogaka, B. O., Bett, H. K., & Ng'ang'a, S. K. (2021). Socioeconomic factors influencing the choice of climate-smart soil practices among farmers in western Kenya. *Journal of Agriculture and Food Research*, *5*, 100168. https://doi.org/10.1016/j.jafr.2021.100168

Mthembu, B. E., Everson, T. M., & Everson, C. S. (2019). Intercropping for enhancement and provisioning of ecosystem services in smallholder, rural farming systems in KwaZulu-Natal Province, South Africa: a review. *Journal of Crop Improvement*, *33*(2), 145–176. https://doi.org/10.1080/15427528.2018.1547806

Prospere, K., McLaren, K., & Wilson, B. (2014). Plant species discrimination in a tropical wetland using in situ hyperspectral data. *Remote Sensing*, *6*(9), 8494–8523. https://doi.org/10.3390/rs6098494

Rebecca. (2020). *Sophisticated Cropping Patterns Mapping in Kenya Using Vegetation Indices and Sophisticated cropping patterns mapping in Kenya using vegetation indices and phenology from Sentinel-2 and Sentinel-1 radar backscatter data. October*. https://doi.org/10.20944/preprints202010.0517.v1

Richard, K., Abdel-Rahman, E. M., Subramanian, S., Nyasani, J. O., Thiel, M., Jozani, H., Borgemeister, C., & Landmann, T. (2017). Maize cropping systems mapping using rapideye observations in agro-ecological landscapes in Kenya. *Sensors (Switzerland)*, *17*(11), 1–16. https://doi.org/10.3390/s17112537

Robbins, O. (2022). *Monocropping: A Disastrous Agricultural System*. https://foodrevolution.org/blog/monocropping-monoculture/

Rusinamhodzi, L., Corbeels, M., Nyamangara, J., & Giller, K. E. (2012). Maize–grain legume intercropping is an attractive option for ecological intensification that reduces climatic risk for smallholder farmers in central Mozambique. *Field Crops Research*, *136*, 12–22. https://doi.org/https://doi.org/10.1016/j.fcr.2012.07.014

Schmidt, K. S., & Skidmore, A. K. (2003). Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, *85*(1), 92–108. https://doi.org/10.1016/S0034-4257(02)00196-7

Shafri, H. Z. M., Anuar, M. I., Seman, I. A., & Noor, N. M. (2011). Spectral discrimination of healthy and ganoderma-infected oil palms from hyperspectral data. *International Journal of Remote Sensing*, *32*(22), 7111–7129. https://doi.org/10.1080/01431161.2010.519003

Shapiro .S.S & Wilk .M.B. (1965). *An analysis of variance test for normality (complete samp1es)*.

Sobhan, I. (2007). Species discrimination from a hyperspectral perspective. In *ITC Dissertation* (Vol. 150). http://www.itc.nl/library/papers_2007/phd/sobhan.pdf

Stomph, T. J., Dordas, C., Baranger, A., de Rijk, J., Dong, B., Evers, J., Gu, C., Li, L., Simon, J., Jensen, E. S., Wang, Q., Wang, Y., Wang, Z., Xu, H., Zhang, C., Zhang, L., Zhang, W. P., Bedoussac, L., & van der Werf, W. (2020). Designing intercrops for high yield, yield stability and efficient use of resources: Are there principles? In *Advances in Agronomy* (1st ed., Vol. 160, Issue 1). Elsevier Inc. https://doi.org/10.1016/bs.agron.2019.10.002

Thenkabail, P. S., Enclona, E. A., Ashton, M. S., & Van Der Meer, B. (2004). Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing of Environment*, *91*(3–4), 354–376. https://doi.org/10.1016/j.rse.2004.03.013

Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., & Skidmore, A. K. (2005). Tropical mangrove species discrimination using hyperspectral data: A laboratory study. *Estuarine, Coastal and Shelf Science*, *65*(1–2), 371–379. https://doi.org/10.1016/j.ecss.2005.06.014

Zhao, J., Fang, Y., Chu, G., Yan, H., Hu, L., & Huang, L. (2020). Identification of leaf-scale wheat powdery mildew (Blumeria Graminis F. sp. tritici) combining hyperspectral imaging and an SVM classifier. *Plants*, *9*(8), 1–13. https://doi.org/10.3390/plants9080936

APPENDICES

Appendix A

UNIVERSITY OF TWENTE RESEARCH
QUESTIONNAIRE FOR CROP MANAGEMENT PRACTICES IN BUSIA COUNTY

This survey aims to obtain information related to maize crop management including the maize cropping patterns. We will ask you questions about one or more of your field plots and we would like to visit those plots with you after the questions. The plots are/should have been planted with maize and/or intercropped maize or other sole crop(s) e.g. groundnut in one plot.
Do you CONSENT to have this information asked? Y ☐ *N* ☐

1	Date and time		
2	Location of farmer's household	Village:	X:
		Ward:	Y:
		Crop field ID:	
3	How many plots do you have?	__plots	

A. Farmer interview sheet [one sheet per plot, maximum of three plots per farmer]

0	Crop field ID				
1	What is the size of the plot (ha)				
2	How many crops were grown between March 2022 and July/Aug 2022 (Long rain)				
	and Sept 2022-Feb 2023 (Short rain)?				
3	Questions (ask them crop by crop)	1st crop	2nd crop	3rd crop	Notes/codes
4	Which crop(s)?				Can be Maize monocrop (M), Maize intercrop (iM), Cassava (C), Groundnut (G),
5	If intercropped, what is the dominant crop?				
6	Type of intercropping				Can be Mixed (M), strip (S), row (R)
7	What was the method of planting				Manual or mechanical

8	Date of land preparation (clearing)				Month and week (1, 2, 3, 4)
9	Date of planting of crop(s)				Month and week (1, 2, 3, 4)
10	Date of flowering				Month and week (1, 2, 3, 4)
11	Date of harvest (or expected harvest date)				Month and week (1, 2, 3, 4)
12	What factors influence the choice of cropping pattern that you practise?				
13	Challenges experienced (is the crop healthy or malnourished?)				
	Notes:				

2021 Data

1	How many crops were grown March 2021 and July/Aug 2021 (Long rain)				
	and Sept 2021-Feb 2021 (Short rain)?				
2	Questions (ask them crop by crop)	1st crop	2nd cro p	3rd cro p	Notes/codes
3	Which crop(s)?				Can be Maize monocrop (M), Maize intercrop (iM), Cassava (C), Groundnut (G),
4	If intercropped, what is the dominant crop?				
5	Type of intercropping				Can be Mixed (M), strip (S), row (R)
6	What was the method of planting				Manual or mechanical
7	Date of land preparation (clearing)				Month and week (1, 2, 3, 4)
8	Date of planting of crop(s)				Month and week (1, 2, 3, 4)
9	Date of flowering				Month and week (1, 2, 3, 4)
10	Date of harvest (or expected harvest date)				Month and week (1, 2, 3, 4)
11	What factors influence the choice of cropping pattern that you practise?				
12	Challenges experienced				

B. Plot data sheet [one sheet per plot, maximum of three plots per farmer]

0	Field boundary ID		
1	Date and time		
	Measurements		
2	Corner Coordinates	X1: X2: X3: X4:	Y1: Y2: Y3: Y4:
3	Field length and width (m)	L:	W:
4	Field size (ha)	field measurement:_____ha	
5	Soil condition	Dry/Wet/Flooding with …... cm water level	
6	Plant height (cm), 3 reps	(a) (b)____(c)________(d) average:___cm	
7	Maize plant age	________days	
8	*Take photos of the field and the surrounding area (N, E, S and W). Draw sketch facing to the north*		

Notes

Appendix B

Representative results of individual bands *p*-values from the Mann-Whitney *U* test.

Nonparametric Tests

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The distribution of w1 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.030	Reject the null hypothesis.
2	The distribution of w2 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.022	Reject the null hypothesis.
3	The distribution of w3 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
4	The distribution of w4 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
5	The distribution of w5 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.019	Reject the null hypothesis.
6	The distribution of w6 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
7	The distribution of w7 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
8	The distribution of w8 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.027	Reject the null hypothesis.
9	The distribution of w9 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.030	Reject the null hypothesis.
10	The distribution of w10 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.022	Reject the null hypothesis.
11	The distribution of w11 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.029	Reject the null hypothesis.
12	The distribution of w12 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.024	Reject the null hypothesis.
13	The distribution of w13 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.022	Reject the null hypothesis.
14	The distribution of w14 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.022	Reject the null hypothesis.
15	The distribution of w15 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.018	Reject the null hypothesis.

60	The distribution of w60 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.110	Retain the null hypothesis.
61	The distribution of w61 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.121	Retain the null hypothesis.
62	The distribution of w62 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.140	Retain the null hypothesis.
63	The distribution of w63 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.136	Retain the null hypothesis.
64	The distribution of w64 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.157	Retain the null hypothesis.
65	The distribution of w65 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.153	Retain the null hypothesis.
66	The distribution of w66 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.166	Retain the null hypothesis.
67	The distribution of w67 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.161	Retain the null hypothesis.
68	The distribution of w68 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.193	Retain the null hypothesis.
69	The distribution of w69 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.171	Retain the null hypothesis.
70	The distribution of w70 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.153	Retain the null hypothesis.
71	The distribution of w71 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.166	Retain the null hypothesis.
72	The distribution of w72 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.161	Retain the null hypothesis.

216	The distribution of w216 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.050	Retain the null hypothesis.
217	The distribution of w217 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.052	Retain the null hypothesis.
218	The distribution of w218 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.035	Reject the null hypothesis.
219	The distribution of w219 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.036	Reject the null hypothesis.
220	The distribution of w220 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.028	Reject the null hypothesis.
221	The distribution of w221 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.024	Reject the null hypothesis.
222	The distribution of w222 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
223	The distribution of w223 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.017	Reject the null hypothesis.
224	The distribution of w224 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.014	Reject the null hypothesis.
225	The distribution of w225 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
226	The distribution of w226 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
227	The distribution of w227 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.032	Reject the null hypothesis.
228	The distribution of w228 is the same across categories of Pattern.	Independent-Samples Mann-Whitney U Test	.035	Reject the null hypothesis.

a. The significance level is .050.

b. Asymptotic significance is displayed.

Appendix C

Classification report as obtained in RF model.

```
# Generate classification report
report = classification_report(y_test, y_pred, target_names=['Maize', 'iMaize'], zero_division=1)
print(report)
```

✓ 0.1s

```
              precision    recall  f1-score   support

       Maize       0.91      0.71      0.80        14
      iMaize       0.50      0.80      0.62         5

    accuracy                           0.74        19
   macro avg       0.70      0.76      0.71        19
weighted avg       0.80      0.74      0.75        19
```