CHARACTERIZATION OF SORGHUM AND GREEN GRAM FOR DATA ESTIMATION USING EARTH OBSERVATION IN MACHAKOS AND THARAKA-NITHI COUNTIES, KENYA

MANZI HILDA KALEKYE

A99/27646/2019

A Thesis Submitted in Fulfillment of the Requirements for the Award of the Degree of Doctor of Philosophy in Agronomy, School of Agriculture and Environment Sciences

February, 2024

DECLARATION

This Thesis is my original work and has not been presented for the award of a degree in any other university or any other award

Signature …… Date …….

Hilda K Manzi

A99/27646/2019

Supervisors

We confirm that the work reported in this thesis was carried out by the candidate under our supervision and has been submitted with our approval as university supervisors

Prof. Joseph Gweyi Onyango

Department of Agricultural Science and Technology

Kenyatta University

Signature ---------- Date ----------

Dr. Shadrack Ngene

Kenya Wild Services

Signature---------- Date ----------

DEDICATION

To my beloved husband Titus Mwaya and our children: Kyla Imani Mwaya and Adrian Muli Mwaya.

ACKNOWLEDGEMENTS

I thank the Almighty God for his present help, without which this work would not have been accomplished.

I thank my supervisors Prof. Joseph Onyango Gweyi and Dr. Shadrack Ngene for their unfailing guidance during proposal development, field, data analysis, and thesis writing. It was a pleasure and privilege working under their very supportive and strict guidance. I am indebted to the farmers of Katangi-Ikombe ward and Chiakariga ward together with the extension officers for all the support they offered during field research work. I am indebted to Prof. James Owouche and Dr. Kevin Obiero Kenya climate-smart project (KSCAP) evaluators who were very encouraging during the research period.

I thank the Kenya Climate Smart Agriculture project through the Ministry of Agriculture, Livestock and Fisheries of the National Government for funding my tuition and research work. I also thank the County Government of Machakos through the Ministry of Agriculture Livestock and Fisheries and the Directorate of Personnel Management who granted me study leave to pursue this course. Finally, I thank my family for all the moral and emotional support they offered me throughout the course of this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION AND ACRONYMS

AEZ	Agro-Ecological Zone
AGB	Above Ground Biomass
APSIM	Agricultural Production Systems Simulator
APAR	Absorbed Photosynthetically Active Radiation
AVHRR	Advanced Very High-Resolution Radiometer images
BPNNOK	Propagation Neural Network-Ordinary Kriging model
BPNN	Propagation Neural Network
CERES	Crop Environment Resource Synthesis Model
CHIRPS	Climate Hazard Group Infrared Precipitation with Station Data
CM SAF	Satellite Application Facility on Climate Monitoring
DEM	Digital Elevation Model
DSSAT	Decision Support System for Agrotechnology Transfer
DSM	Digital Surface Model
DTM	Digital Terrain Model
EC	European Commission
EPIC	Environmental Policy Integrated Climate
EOS AM	Terra Earth Observation System Automated Messaging
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
ETM+	Enhanced Thematic Mapper Plus
FAO	Food and Agriculture Organization
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FEWSNET	Famine Early Warning System Network
FOV	Field of View

GIEWS	Global Information and Early Warning System
GIS	Geographical Information System
GPP	Gross Primary productivity
GPS	Global Positioning Systems
IFOV	Instantaneous-Field-of View
ITCZ	Inter Tropical Convergence Zone
LAD	Leaf Angle Distribution
LAI	Leaf Area Index
LST	Land Surface Temperature
MAM	March April May
MARS	Monitoring Agriculture with Remote Sensing Systems
MATLAB	Matrix Laboratory
MODIS	Moderate Resolution Imaging Spectroradiometer
NPP	Net Primary Productivity
NDVI	Normalized Difference Vegetation Index
NIR	Near-infrared
NOAA	National Oceanic and Atmospheric Administration
OND	October November December
OODBMS	Object-Oriented Database Management System
PAR	Photosynthetically Active Radiation
PC	Principal Components
SPECCHIO	Standard Plant Species Spectral Online database
PH	Plant Height
RENDVI	Red Edge Normalized Difference Vegetation Index

RDBMS	Relational Database Management System
SID	Sun Direct Irradiance
SIS	Sun Irradiance
SMC	Soil Moisture Content
SMI	Soil Moisture Index
SOM	Soil Organic Matter
SOTWIS	Soil Terrain World Information System
SSI	Soil Salinity Index
SWIR	Shortwave Infrared
TM	Thematic Mapper
THVC	Traditional High-Value Crop
TRMM	Tropical Rainfall Measuring Mission
USAID	United States Agency for International Development
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VI	Vegetation Index

ABSTRACT

Data management in agriculture is very vital for the planning of any country, including Kenya. A study was conducted to establish the role that remote sensing data can play in the assessment of spatial variability in crop data and the provision of crop farm information. The main objective of the study was to develop a remote sensing-based crop data information system for estimating green gram and sorghum data, respectively, under farm field conditions. The study was carried out in Ikombe-Katanga area of Machakos County for the green gram crop and in Tharaka Nithi area for the sorghum crop. Estimating crop data was based on three approaches: crop area estimation, spectral signature library development for sorghum and green gram, and assessment of microclimates within agroecological zones. The following parameters were estimated and used to calculate the most important crop data for green gram and sorghum for the October, November, and December rains. The parameters selected for the development of this estimation model were vegetation indices, biomass, leaf area index (LAI), enhanced vegetation index (EVI), soil moisture index, soil organic carbon, rainfall, land surface temperature, soil pH, soil nutrients (NPK), and evapotranspiration. The first crop data estimation involved crop area, determined by assessing sorghum and green gram cropping pattern data. The identified cropping pattern for the green gram study area and crop area estimates were mixed crop (maize and green gram), mixed crop (maize and beans), green gram, mixed crop (maize and pigeon pea), mixed crop (maize and cowpea), and maize with 2445.93 ha, 10,034 ha, 5981 ha, 4697.82 ha, 3743.82 ha, and 586.35 ha, respectively. The sorghum crop patterns were mixed crop (sorghum and beans), mixed crop (sorghum and cowpea), and sorghum, with crop area estimates of 1988.46 ha, 961.65 ha, and 469.62 ha, respectively. Furthermore, the development of spectral signature libraries was done, and the spectral reflectance ranges for sorghum and green gram were 0.230064 to 0.321126 and 0.26900 to 0.07466 across all bands, respectively. The results further revealed the existence of microclimates within agroecological zones IV and V and four microclimatic zones within the agroecological zones lower midland IV and V of Tharaka Nithi and Machakos counties. Using data for October, November, and December (OND) rains and cropping season data, it was possible to estimate crop yield. Twelve parameters were analyzed and ran through a random forest machine learning algorithm to generate sorghum yield estimates. The validation of the model was carried out using root mean square error (RMSE) and root mean absolute error (RMAE), with results showing that RMSE was 4.036 with R^2 of 0.98 and RMAE was 3.022 for the green gram crop, while for sorghum, RMSE was 6.51 with R^2 of 0.99 and RMAE of 5.5. The yield estimates were 4.5 bags/acre for green gram and 9 bags/acre for sorghum, respectively. The data estimation under farm field conditions was sufficiently optimized using the farm crop data estimation (FCropDesti) tool developed from this work using ArcGIS software. The study also confirms that employed methodologies were important in creating a homogenous environment on farms for the identification of cropping patterns, which further determine crop data such as crop area, crop type, and crop yield. Crop data estimation is important for policymakers at the county and country level to implement climate-smart agriculture. The research can be replicated in other important crops in the country.

CHAPTER ONE: INTRODUCTION

1.1 Background information

Statistics on crop production are crucial for agricultural planning and policymaking and this key in enables planning within agriculture (FAO, 2023). Sustainable agricultural production requires robust information and data acquisition that that can benefit from an expanding body of knowledge. Gaining knowledge by integrating data from multiple sources can yield significant and practical insights. At the farm level, contemporary farmers are confronted with an extensive array of data on a daily basis in order to make decisions that affect their livelihoods. Factors such as financial resources and loan accessibility, soil health, weather conditions, irrigation practices, market conditions, early-warning systems for diseases and pests, government-related information, and subsidies all contribute to the decision-making process at the farm level (FAO and ITU, 2019). Sub-Saharan Africa has been faced with challenge of low productivity level emanating from not only poor soil but also a lack of information on the amount of essential inputs required to raise productivity (Thomas, 2020). This is because there is a lack of adequate crop production statistics that can inform a turnaround in productivity.

In Kenya, where smallholder farmers depend on agriculture for income, crop production statistics are even more important (Amwata, 2020). The importance of agriculture in Kenya and its role in shaping the economy of the country through employment, food security, and national self-reliance cannot be overemphasized (Birch, 2018). Kenya reports crop area and yield statistics using different methods based on state institutional mechanisms. The crop production statistics are at different small scales, calculated for the regional scale (Birch, 2018). These statistics are often available at the district scale and, to some extent, at the sub-district scale. Sibanda and Mwamakamba (2021) indicate that the district/sub-district scale is adequate for the planning of agricultural inputs where the majority of these transactions happen, as well as for setting up the minimum support price and for national/international selling and procurement decisions. For the purpose of agricultural statistics at the regional and national levels, as well as for quantifying

the output and productivity of agricultural systems at the research plot level, numerous techniques have been developed (Sapkota *et al*., 2016). However, a revolution in the agriculture system is likely to pose a challenge in data estimation because different systems require different technologies (Sapkota *et al*., 2016). The reliability and timeliness of agricultural production data are insufficient to meet crucial information needed in many African countries (Murphy *et al*., 1991). For example, Sapkota *et al*. (2016) indicate that crop cutting is prone to bias and errors if applied to a crop with raised beds. Precise data estimation in smallholder farms has remained a challenge due to the heterogeneous nature of crop performance, intercropping, specifically relating but not limited to mixed cropping, and continuous planting occasioned by farmer needs (Sapkota *et al*., 2016).

Monke *et al*. (2019) indicate that the changing market in economies demands reliable agricultural information for quick decision-making in the agriculture sector, which is currently a challenge. The demand for high-quality agricultural statistics for evidence-based decision-making, planning, and the development of policies for adaptation to factors such as climate change are quite pertinent. The need for timely information to increase agricultural outputs, raise productivity, and adapt to climate change creates a demand for solutions towards adequate and reliable data in the agriculture sector. A successful farming system in developed countries has been well organized through a farmer information system (Verheye *et al*., 2011). Study by Panigrahy *et al*. (2011) shows that the use of such systems is important for the organization and ease of management of farmers through groups and extension channels for he purpose of improving production. The availability of smallholder systems among smallholders is still in its formative stages, even though the demand is high (Karienye *et al*., 2019; Srinivasarao *et al*., 2013). In general, monitoring systems such as the use of remote sensing data for yield estimation to secure compliance with restrictions and standards in terms of specific production guidelines have become necessary (Fritz *et al*., 2019). According to Jones *et al*. (2017), data scarcity and limitation are key challenges in agricultural monitoring and forecasting. The lack or limited agricultural monitoring systems at regional and national levels has been attributed to little or no information sharing (Fritz *et al*., 2019).

The variability of crop data and related variables, as well as natural resource use, as documented by Ngure *et al*. (2021), are mainly due to the short- and long-term variation of weather and climate conditions. The results from models, according to Makowski *et al*. (2006), can be used to make appropriate decisions and to provide farmers with alternative options for their farming system.

A lot of crop models used for farm data, and specifically yield, are limited when it comes to data estimation under such conditions. Gaining an understanding of these factors, how they interact with the environment and how they affect management, and expanding model capabilities is a challenge for agricultural scientists (Boote *et al*., 1996). This research integrates crop models with spatial data to overcome the challenge of incorporating all necessary factors for field data estimation under current models. Remote sensing emerges as an important technology (Qi *et al*., 2000) for the investigation of spatial variability in crop yield and the provision of crop farm data. In the recent past, aerial images have also been used in yield estimation in precision farming (Rembold *et al*., 2015). The images of the crop spectral characterized by Reynolds *et al*. (2000) and Schmedtmann and Campagnolo (2015) inform the importance of remote sensing in yield prediction. Vegetation analysis and monitoring of changes in vegetation patterns have been found to play a crucial role in providing data on crop vigor (Shammi and Meng, 2021).

Estimating crop yield before the harvest is very important in agriculture since variation in crop yield across different years impacts trade, market prices, and the overall food supply (Hayes and Decker, 1998). At field level, crop yield data inform farmers' timely decisions for a particular situation, such as whether or not to grow alternative crops or avoid farming in a particular season (Hayes and Decker, 1998). Various estimation models have been used previously in yield estimation; e.g., Baez-Gonzalez *et al*. (2005) who used Landsat ETM+ (enhanced thematic mapper) data in an NDVI model to predict the yield of paned corn in Sinaloa, Mexico. In this research, an average error of 9.2% in the estimation of corn yields was reported. Yang *et al*. (2008) used the United States Department of Agriculture's (USDA) EPIC model for predicting yield in some parts of China. In their investigation, it was evident that the

statistical and simulated crop yield was under 10%. Baez-Gonzalez *et al*. (2005) modeled corn yield in Mexico with NDVI derived from NOAA-Advanced Very High-Resolution Radiometer (AVHRR) images. In this investigation, 89% and 76% variability in yields were found for the crops grown under irrigation and non-irrigated conditions, respectively. Remote sensing and optical satellite systems have been used in the development of national crop inventory programs (Husak *et al*., 2008). They have been found to provide important data for the estimation of cultivated area and yield for commercially important crops.

In this research, green gram (V. radiata L.), or mung bean, was the crop investigated. It has recently become an important legume among smallholder farmers in Kenya, and consequently, the consumption of mungbean is increasing. According to Hargrave (2007), Hill (1987), Purseglove (2003), and Samant (2014), the mung bean is a very important short-season crop in the world. In Kenya, green gram is grown successfully in the semi-arid areas of Machakos, Kitui, Tharaka-Nithi, and Makueni counties due to its ability to mature early (USAID, 2013). Machakos County, which is the study area, leads in green gram production. Green gram survives under different soil and climatic situations due to its drought tolerating abilities (Wambua *et al*., 2017). Furthermore, the crop is adapted to a variety of soil conditions, including poor soils, because it forms associations with Mychorrhiza, and hence plays a role in both environmental conservation and food security (Wambua *et al*., 2017). According to Singh *et al.* (2002), smallholder farms supply up to 90% of the food in developing nations, making them the primary suppliers of food in these regions. Understanding the production of such vital crops is important for planning and monitoring food security (Mugo *et al*., 2016). Green gram has been recommended in various government projects in Kenya as one of the adaptation strategies for climate change (Mugo *et al*., 2020). They consider green gram to be one of the main crops in semi-arid areas of Kenya. Optimization of green gram crop productivity is very essential to addressing food security in Kenya and Africa at large. In Mugo *et al.* (2016) study, the growing condition of green gram were modelled for appropriateness under current and future conditions. The conclusion of the study was that suitable green gram-growing areas will increase during October, November, and December (OND) and

decrease during March, April, and May (MAM). The high temperature as a result of climate change places green gram, which performs well in dry conditions, as a better crop for the dry regions (Mugo *et al*., 2016). Planting green grams in suitable areas is likely to solve the food security problem by increasing food production (Mugo *et al*., 2020). A number of challenges are facing green gram production, including low technology adoption, poor storage, an inadequate policy framework, and limited value-added opportunities. This is exacerbated by the lack of adequate crop data that could conveniently shape a proper policy framework for the green gram crop in Kenya if such data were available.

Sorghum, on the other hand, is one of the most important food crops in sub-Saharan Africa. Production of food in Kenya over the years has been declining (FAO, 2013). The declines in crop productivity, especially maize, have been linked to unreliable rainfall and frequent drought episodes, inadequate fertilizers or unbalanced believed fertilizers and gas emission believed to be related to climate change events (Meehl and Stocker, 2007; Ntinyari and Gweyi-Onyango 2021; Ntinyari *et al.,* 2023). This has led to the introduction of crops tolerant to drought, coupled with high water use efficiency that can survive under changing weather patterns (Twomlow *et al*., 2008). Crops such as sorghum with high water use efficiency have been introduced to semi-arid areas such as Kitui, toand the records show than on average the can yield as high as 1.57 metric tons (FAO, 2013). Further, FAO (2013) indicates that sorghum is tolerant drought well, making it an important crop for the arid and semi-arid lands (ASALs) that are highly susceptible to adverse weather patterns. The popularity of sorghum in Kenya is high, driven by its adaptability to different climates, coupled with its industrial uses in alcohol production. The paper presents the shift from semi-arid to lower highlands, where the potential for growing sorghum is high but low for the maize crop. It is a cereal crop that is quantitatively ranked the fifth most important cereal grain in the world after wheat, maize, rice, and barley (Kigen *et al*., 2014). Africa is one of the largest sorghum producers, accounting for one-third of global production (Taylor *et al*., 2003). The suitability and adaptability to tropical conditions prevalent in Africa explain the dominance of sorghum (Mwema and Mulinge *et al*., 2013). In sub-Saharan Africa (SSA), the crop is prominently used as

a viable cereal crop for the most food-insecure households (Muui *et al*., 2013). Due to increasing trends in global warming and climate change, sorghum is a promising alternative for improved food and income security, compared to other staples such as maize that often fail due to drought (Muui *et al*., 2013). In Kenya, sorghum production is concentrated in marginal and semi-arid areas whose rainfall regimes are low and erratic and where the temperatures are high (Mwema and Mulinge 2013). Currently, proper data on the sorghum crop is inadequate, and this has affected its adoption by farmers due to lack of data on the productivity and profitability of sorghum farming systems.

1.2 Statement of the problem

Effective crop data estimation has been the biggest challenge in Kenya. This has affected the ability to plan food security. Husak *et al*. (2008) indicate that overestimating the crop area is responsible for a high level of undernourishment in Africa. Many times, the planted area does not translate into cropped areas. Changes in dynamics and heterogeneous farming activities among these smallholder farmers make it even more difficult to collect ground data, which can sometimes be laborious (FAO, 2011). Mapping crops is an important measure to determine the amount of food produced, where it is produced, and when it is produced. The increasing human population and decreasing food availability push for the need to monitor crop production (Dixon *et al*., 1994).

Farming in semi-arid regions such as Machakos and Tharaka Nithi Counties has expanded over the years as the demand for food increases with dwindling land sizes in high-potential areas. Sorghum and green gram are some of the climate-smart crops whose uptake by farmers is very low. To increase uptake of these crops among smallholder farmers in semi-arid regions, crop data that guides the production trends and informs both farmers and the government's practices, management and planning are crucial. To develop policies that address crop data estimation for accurate planning, there is a need to recognize the role that crop models play in understanding the impact of climate change on agricultural production and overall food security (Makowski *et al*., 2006).

In this study, a stepwise crop data estimation tool was developed with a focus on ensuring its applicability under farm field conditions. The first step involved digitizing the farm boundaries to increase accuracy in crop area estimation and the development of spectral signatures for crop identification. Second, the agroecological zones were examined for microclimatic zones to ensure a homogenous environment for the estimation of yield. Lastly, crop yield estimation was done using ten different remote-sensing parameters, and the crop yield for sorghum and green gram was estimated and validated using field data. This is the first that such sequential and structured has been reported in Kenya and SSA region.

1.3 Objectives

The main objective of the study was to develop an earth observation-based crop data estimation tool for green gram and sorghum under farm field conditions in Machakos and Tharaka Nithi Counties.

1.3.1 The specific objectives were

i. To establish the spatial distribution and spectral signatures of sorghum and green grams at farm field conditions for crop type and crop area estimation
ii. To establish micro-climatic zones within the existing agroecological zone IV and V of Ikombe katangi and Chiakariga study areas
iii. To estimate and model yields of Sorghum and green grams at farm field conditions
iv. To develop an earth observation-based farm data estimation tool for sorghum and green grams

1.4. Research Hypothesis

H_0 Spatial distribution and spectral signatures of sorghum and green grams cannot be determined under farm field conditions for crop type and crop area

H_0 Existing agro-ecological zones IV and V of Ikombe katangi and Chiakariga the study area in Kenya have no clear micro-climatic zones

H_0 Sorghum and green gram crop yield cannot be estimated under farm field condition.

H_0 Earth observation-based farm data estimation tool cannot be developed for sorghum and green gram

1.5 Justification of the study

The growing of crops has become a big challenge for small-holder farmers, as they are unable to monitor the crop in the field. The challenge is intensified by agriculture's extreme susceptibility to changing weather patterns. The negative impacts come as a result of increased temperatures, unpredictable weather patterns, shifting boundaries of the agroecosystem, emerging pests and diseases, and more frequent unpredictable weather events. In most countries, especially Kenya, the assessment of food security is further challenged by a lack of crop data that can be delivered timely and whenever needed.

The continuous availability of spatial data in sorghum and green grams in a timely and effective manner will provide the prerequisite data that the government and research institutions will require to address food production needs in Kenya. This information is essential for improving crop yield prediction, management, proper resource utilization, economic planning, and monitoring development interventions. Forecasting and monitoring agricultural production is very important for local food demand and overall food security.

Farm data estimation in a timely manner is critical for addressing climate change issues as well as solving food security problems that are exacerbated every day with a growing population. Working with existing information, technologies, and principles of sustainable agriculture, farm data estimation plays a key role in ensuring minimized food losses and the sustainable maximization of available resources for agricultural production. This will ensure smallholder farmers can solve all crop production problems for sorghum and green gram, thereby averting losses and enabling increased income at the household level. The financing and planning of policymakers in Kenya cannot be accurate without accurate farm data.

1.5 Significance of the study

The study addresses the effect of climate variability on crop and food security issues by providing a crop data estimation tool that will enable the management of crops under farm field conditions. This crop data information will provide real-time spatial data for use in monitoring sorghum and green gram crops. The earth observation-based farm data estimation tool enables monitoring of crop data such as crop health, crop area, and crop yield at farm field conditions. This information will be useful to farmers, agricultural researchers, agriculture-promoting institutions and non-governmental organizations such as the World Bank and Food and Agriculture Organization (FAO), agricultural policy makers (including the Ministry of Agriculture, Livestock, and Fisheries), extension officers, and farmers, especially those in Sub-Saharan Africa.

The technology has been developed to integrate farmer practices while allowing crop monitoring for early decision-making. The study attempts to solve the problem of crop area estimation through digitization of farm boundaries, upon which crop classification on specific cropped areas improved the accuracy in estimating the area that the crop has been grown in any particular season. The study also creates crop spectral signatures for sorghum and green gram using multitemporal datasets. This lets the researchers identify the crops, which makes area estimates more accurate. This technology can be applied to develop spectral crop libraries in Africa for increased precision in crop identification that determines the crop area as well as enabling crop monitoring.

The study further enables the assessment of agro-ecological zones for any existing microclimatic zones in an attempt to create a homogeneous environment from which crop yields and important crop data can be estimated. The models created are assembled in an ARCGIS tool and can be accessed for decision-making relating to food security in Kenya and Africa at large. The research provides evidence and the possibility of creating crop spectral libraries that would increase the use of remote sensing data in crop data estimation, which is key for policymakers.

1.6 Conceptual framework of the study

In this study, the model of the research uses the integrated approach schematized in Figure 1.0. The method involves developing spatial data outputs in steps, such as farm delineation, spectral signature tools, and agro-ecological sub-zonation tools, from objectives 1 to 2. These steps create a consistent environment for crop yield models to work correctly in farm field conditions, which is objective 3.Further integration of three models provides the final aspect of the concept, which is the crop data estimation tool, which is objective 4. The data inputs for the approach include farm boundary data, crop spectral reflectance data, agroecological zone data, crop data, soil data, and climate data.

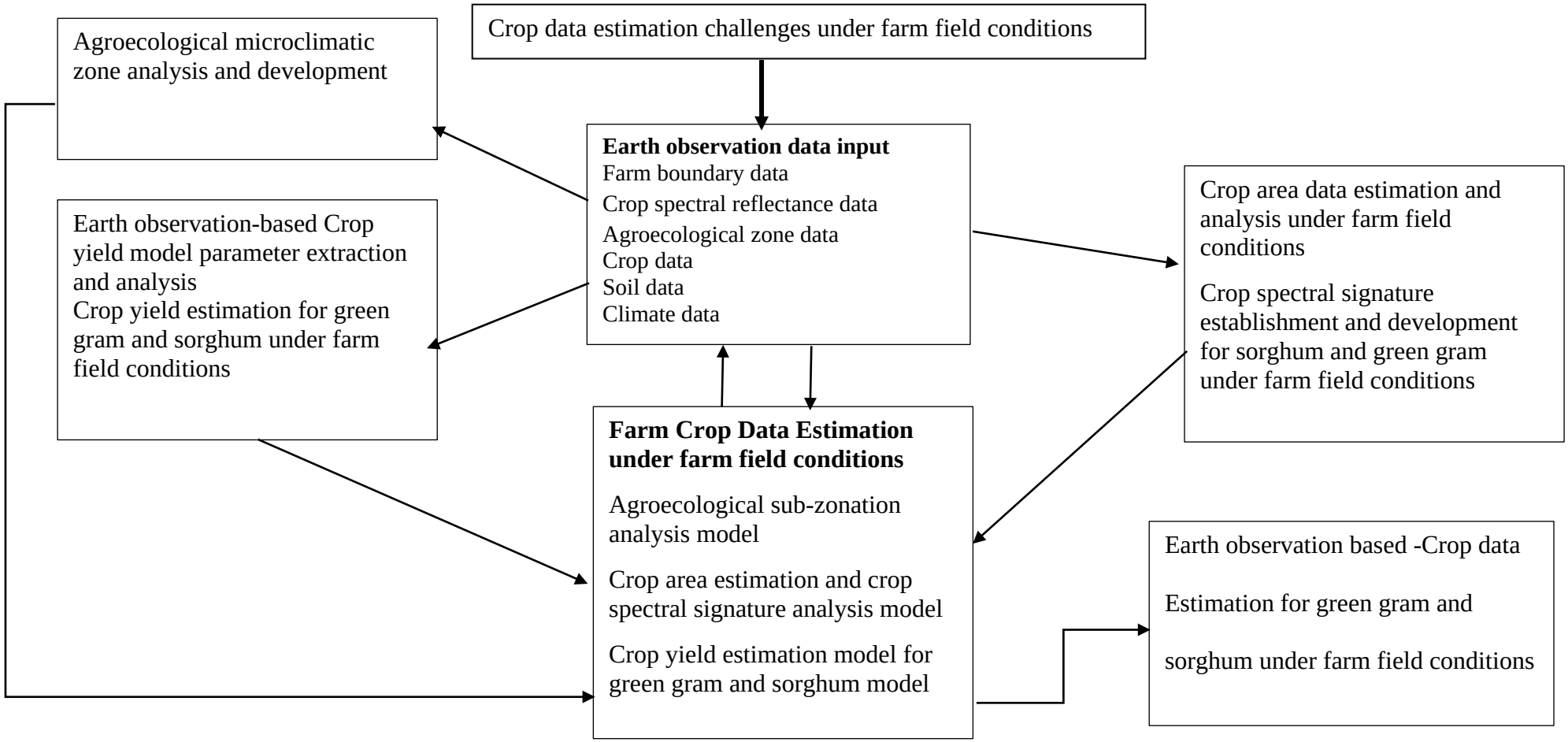


Figure 1.1 Conceptual framework

CHAPTER TWO: LITERATURE REVIEW

2.1 Agricultural production

Since agriculture provides the majority of household income in developing nations, UNEP (2011) has demonstrated the role that agriculture plays in their economic and social development. More than 60% of people living in Africa rely on agriculture as a source of livelihood (FAO, 2004). In South Asia, 75% of the poor living in rural areas depend on rain-fed agriculture, livestock, and forests for subsistence (Sapkota *et al*., 2015). Waongo *et al*. (2015) reported that high rainfall variability and the lack of adaptation strategies limit production in smallholder farming systems across Sub-Saharan Africa. In the end, it affects food security and the economy, leading to more poor people on the African continent. Climate change, without adaptation, is projected to negatively impact crop productivity. Indeed, temperature increases of 2^{o}C or more are expected to negatively affect the major crops (i.e., wheat, rice, and maize) in temperate and tropical regions (Meehl and Stocker, 2014). For temperature, Lobell and Gourdji (2012) found that each degree above 30^{o}C reduces crop yield.

Eighty percent of Kenya is made up of dry and semi-arid regions that are extremely vulnerable to the effects of climate change (Mugo et al., 2020 (Manzi and Gweyi-Onyango 2020). These areas are part of the marginal areas of Kenya, highly vulnerable to degradation due to sensitive landscapes. The semi-arid areas of Kenya are characterized by fluctuating rainfall patterns that are highly unreliable. The amount of rainfall is low with very high temperatures, which cause high evapotranspiration rates (Mugo *et al*., 2020). Climate variability in both spatial and temporal forms is one of the causes of decreased agricultural productivity in these areas (Samwel *et al*., 2021). There is general agreement that climate change is changing and that the agricultural sector, among others, will be affected by future climates. Adverse weather events may have an impact on Kenyan agriculture and farmer adaptation, making current agricultural use unsustainable (Samwel *et al*., 2021). The Inter-Tropical Convergence Zone (ITCZ), a low-pressure zone abroad, describes the climatic condition in Kenya (Hills, 1979). The ITCZ occurs as a result

of the intersection of the northeast and southeast trade winds of the two hemispheres (Hills, 1979).

Temperatures can be described as equatorial, characterized by minimal average monthly and annual temperatures closely correlated with altitude (Ngure *et al*., 2021). Bolt *et al*. (2019) describe the temperature (from 1961–2005) for both the short (October, November, and December) and long (March, April, and May) rainy seasons in Kenya as increasing by more than 0.8 °C. They further highlight that the increases have been more significant in the north-eastern and north-western parts of the country in both the short and the long October, November, and December rains. The increases in these particular regions have been documented to be greater than 1°C. The East Africa region where Kenya lies has been documented to be one of the topographically diverse regions and one of those whose meteorological complexity is unique and the only one within Africa (Berhane and Zaitchik, 2014). Berhane and Zaitchik (2014) further indicates that the rainfall variability, i.e., interannual, inter-seasonal, and intra-seasonal time scales, has had a huge impact on rainfed agriculture and food security as a whole. The economy of Kenya is highly dependent on rainfed agriculture. The shortened droughts have had a huge impact on both food security and the economy.

2.2. Agricultural monitoring systems

Agriculture and natural resources are under strong pressure due to population growth, increased consumption of calorie- and meat-intensive diets, and an increasing use of cropland for bioenergy production (Bousbih *et al*., 2019; FAO, 2014; Foley *et al*., 2011; Hills, 1979). Bridging the gap and reducing the pressure on agriculture and natural resources has necessitated various options that include: 1) expansion of land under cultivation; 2) intensification on currently available farmlands; 3) reducing the yield gaps in farmers' fields; 4) raising the yield levels through higher yield varieties; and 5) reducing post-harvest losses and food waste (Charles *et al*., 2010; Keating *et al*., 2014; Struik and Kuyper, 2017). Ramankutty *et al*. (2008) indicated that half of the land suitable for agriculture is already under cultivation today. The negative impacts of this are highlighted by Folley *et al*. (2011) as part of agricultural

expansion and intensification. According to the AGRA Africa Agriculture-Status Report (2014), different climatological scenarios have shown the limitations of diversification options for agropastoral systems that are common in semi-arid regions. Agricultural monitoring on a regional and national level has been in place for decades, as attested by the following examples: the Food and Agriculture Organization's (FAO) Global Information and Early Warning System (GIEWS), the United States Agency for International Development's (USAID) Famine Early Warning Systems Network (FEWSNET), CropWatch in China, and the European Commission's (EC) Monitoring Agriculture with Remote Sensing System (MARS). These systems have tended to function rather autonomously, with little information sharing, with the emphasis on either food security for developing countries or food production for the global market (Fritz *et al*., 2019).

Data recorded by remote sensing satellites mainly assists with the assessment of crop conditions and crop anomalies, which can then be used to infer information on yield, area, and production reductions. However, this approach is not able to provide quantitative crop area and production forecasts, which are ideally needed for food security interventions. Crop growth and yield forecasting models are data intensive and are currently applied only for the US by the USDA system and for Europe (Fritz *et al*., 2019). Moreover, (Fritz *et al*., 2019) puts its that remote sensing-based methods for agricultural statistical forecasting needs historical archives of high-quality statistics, which are not available in all countries.

Remote sensing data can provide timely, synoptic, cost-effective, and repetitive information about the status of the Earth’s surface (Singh *et al*., 2002). Data from remote sensing satellites can provide two components of crop production; yield (Hoefsloot *et al*., 2012; Rembold *et al*., 2013) and acreage(Atzberger, 2013; Zhang *et al*., 2013). Additionally, remote sensing data can provide crop phenological information (Sakamoto *et al*., 2005), stress situations (Hoefsloot *et al*., 2012). Among other things, the retrieved information allows decision makers to understand climatic events and to obtain good information for areas risk assessment (Zhang *et al*., 2013).

Remote sensing remains key in monitoring of agricultural activities since the sector faces problems that are not common to other economic sectors (FAO, 2011).
A study reported by Bobade *et al*. (2010) in the Seoni district of Madhya Pradesh, India, on land evaluation for agricultural planning indicated the incorporation of GIS technology in soil survey data collection. Soil-based GIS data was used for land use suitability and fertility assessment for the Seoni district. They found the use of geospatial technology to be of great importance in farm production planning and decision support. Currently, geospatial information technologies are becoming increasingly important in the decision-making process of land management planning (Singha and Swain, 2016). GIS, coupled with satellite data, gives decision-makers an opportunity to assess landscapes for improved natural resource management. The use of geo-spatial information links local knowledge and science together with national development strategies.

One general benefit of geospatial data is the increased accuracy of data collection and analysis (Atzmanstorfer and Blaschke, 2013). These technologies have been found to successfully enable the management of land-related resources in industrialized countries. The increasing availability of remote sensing images makes it possible to develop monitoring systems capable of automatically producing and regularly updating land cover maps of the considered site (Hoefsloot *et al*., 2012). Remote sensing can significantly contribute to the monitoring of the agricultural sector, as it enables the gathering of information over large areas with a high frequency of revisits (Atzberger, 2013).

2.3 Crop area estimation and crop mapping using remote sensing technology

Smallholder farms are subdivided into several parcels whose boundaries are hard to distinguish (Fritz *et al*., 2008). Spatial information on agricultural fields and how they vary across the region is important in monitoring crop condition and yield. This information is essential for improving crop yield prediction, management, proper resource utilization, economic planning, and monitoring development interventions

(Singh *et al*., 2002). Forecasting and monitoring agricultural production is very important for local food demand and overall food security (Reynolds *et al*., 2000). Remote sensing tools have been widely used in crop area estimation, crop type identification, and yield estimation (Fritz *et al*., 2019). The standard methods of crop area estimation have relied on techniques such as area statistical frame studies (Tsiligirides, 1998). These are some of the field-based survey studies that can take up a great deal of time and resources. These standard methods require a lot of resources to establish crop areas among small-holder farmers (Singha and Swain, 2016). Earth observation data coupled with good cropping pattern training data have revolutionized crop identification, making it easy to understand the spatial distribution of crops (Atzberger, 2013). This has helped to address food security issues that are common in most semi-arid areas of Africa.

Data from remote sensing has played a significant role in identifying different crops, evaluating crop health, and assessing yields (Nellis *et al*., 2009). Supervised and unsupervised classification has been applied for mapping the geographical distribution of crops and characterization of cropping practices (Boitt *et al*., 2014). Different band ratios, multispectral data, and classification schemes have been used to generate crop information such as crop diversity, field size, crop phenology, and soil conditions (Nellis *et al*., 2008). Hyperspectral remote sensing has played a major role in advancing crop classification (Thenkabail and Wu, 2012). Wang *et al*. (2007), for example, used satellite remote sensing of NDVI for analysis of landscape-level patterns of net primary productivity within the U.S.

This research has used similar methods with an improvement, where additional tools that enable crop area estimation through digitized boundaries and creation of homogenous area through reclassification of agro-ecological zones were realized.

2.4 Agro-ecological zone

The role of the agroecological zone in defining the crop environment has been generated from the agroclimatic zone of Kenya. An agroecological zone is an area that is defined by its relevant agroclimatic factors and differentiated by soil pattern. It provides a framework for the ecological land use potential (FAO, 1996). Generalized agro-ecological zones were established by FAO (1996) that were suited for decision-making in agricultural policy. A more detailed agro-ecological zonation was done by Jiitzold and Kutsch (2000), which provided a more differentiated system showing the probabilities and risks in farming. For example, zone groups based on temperature belts defined according to the maximum temperature limits within which the main crops in Kenya can flourish were: cashew and coconuts for the lowlands; sugar cane and cotton for the lower midlands; Arabica coffee for the upper midlands; tea for the lower highlands; and pyrethrum for the upper highlands (Jiitzold and Kutsch, 2000). The main zones were based on their probability of meeting the temperature and water requirements of the main leading crops, i.e., the climatic yield potential, calculated by computer (FAO, 1996). The zones are roughly parallel with Braun's climatic zones of the Precipitation/Evaporation Index, with a few differences influencing the length and intensity of arid periods. This study intends to further generate micro-agroecological zones to reduce spatial variability in all the data sets that shall be generated (Jiitzold and Kutsch', 2000).

2.5 Crop discrimination using remote sensing data

The crop spectral signature plays a significant role in crop identification, especially the mapping of cropped areas (Awad *et al*., 2019). This provides vital information that can be used to forecast crop yield in any given season of planting. Evaluation of the spectral signature of crops has been very useful in crop identification, especially during phenological cycles and spectral similarity (Duong *et al*., 2014; Kachhwaha, 1983). The spectral signature provides useful information that can be retrieved by analyzing the temporal signatures and directional reflectance properties

of vegetation (Wardlow *et al*., 2007). Reflectance is related to the absorption and transmission of each wavelength, thus representing the status of the plant under ambient or experimental conditions (Garriga *et al*., 2014). Since the early 1980s, crop types have been distinguished using temporal and spectral characteristics (Badhwar, 1982). Time periods with the highest differences between crop types are often previously identified (Kyei-Mensah *et al*., 2019).

Nidamanuri and Zbell (2012) indicate that the comparison of spectral signatures between different authors is a problematic issue due to the many different techniques used for the capture of spectral field data. Furthermore, Awad *et al*. (2019) say that the creation of a database of crop spectral signatures is a complex task owing to the demand for consistent data from seeding to harvesting. There are other spectral signature libraries for agriculture research purposes, such as the SPECCHIO Spectral online database maintained by the Remote Sensing Laboratories in the Department of Geography at the University of Zurich (Hueni *et al*., 2009). This database displays metadata about crop characteristics, date of acquisition, vegetation biophysical parameters, soil characteristics, and other important crop information (Hueni *et al*., 2009). Bojinski *et al*. (2003) indicate that the library has been tested, albeit with the challenges of a non-structured multi-criteria query. In addition, crop-specific information is difficult to get due to the hundreds of records that the search brings.

The limitation in spectral signature data for Africa was a motivation to undertake the current research to develop crop spectral signatures for sorghum and green gram in specific locations in Kenya to improve crop management and help establish precision in agricultural practices, especially among smallholder farmers. Heupel *et al*. (2018) found that previous studies on crop-type classification had differences regarding the method applied in analysis, the number and type of data sets, the study area, as well as the availability of field and training data. There is consequently no consistent crop-type classification approach due to multiple regional conditions and characteristics.

2.6 Crop area estimation using spatial data

According to Husak et al. (2008), a 2003 FAO study demonstrated that a significant portion of Africa's undernourishment can be attributed to either an overestimation or an underestimation of the crop area. Many times, the planted area does not translate into cropped areas. This is attributed to germination effects that may result from various environmental conditions (Reynolds *et al*., 2000). These can be soil- or climatic-related conditions that interfere with crop growth. The estimation of cropped area among small-holder farmers in developing countries has faced the challenge of poor estimation due to unreliable data (FAO, 2013). Changes in dynamics and heterogeneous farming activities among these smallholder farmers make it even more difficult to collect ground data, which can sometimes be laborious (FAO, 2011).

Mapping crops is an important measure to determine the amount of food produced, where it is produced, and when it is produced (Nidamanuri and Zbell, 2012). The increasing human population and decreasing food availability push for the need to monitor crop production (FAO, 2011). The advent of satellite imagery presents an opportunity for the estimation of cropped areas but also for the monitoring of the crop at different stages of growth (Heupel *et al*., 2018). Medium-resolution satellite imagery such as Landsat and Sentinel has, on different occasions, used different methodologies in the assessment of the cropped area under agriculture (Cunha *et al*., 2010). For example, area frame sampling has been applied several times to establish cropped areas. The challenge of this particular methodology is the laborious field survey, while the use of satellite images through supervised classification has resulted in pixel mixing (Jamal-Eddine *et al*., 2018). The results are therefore not properly estimated in the long run. When crop classification is carried out by spectral signature identification, the spatial distribution of the crop can be determined (Hueni *et al*., 2009).

2.7 Crop yield estimation using remote sensing data

Remote sensing data has been widely applied to many research problems and practical applications in agriculture (Ren *et al*., 2012). One such solution is the use of remote sensing data for yield estimation. Compared to traditional data collection methods, the ability of remote sensing techniques to provide timely information over a wide spatial extent at a wide range of spatial, temporal, and spectral resolutions is appreciated by numerous users for yield estimation (Awad, 2019). Agriculture is one of the main users of remote sensing data (Moulin *et al*., 1998) and numerous research efforts have devoted time to seeking relationships between remotely sensed spectral information and crop yields and consequently obtaining robust estimation and forecasting for agricultural productions (Idso *et al*., 1977). There are two strategies that have been used in estimating crop yields based on remote sensing data (Ferencz *et al*., 2004). The first one is based on crop growth models, which incorporate remote sensing data into agrometeorological or biophysiological models (Ferencz *et al*., 2004). For example, Ansarifar *et al*. (2021) highlight some of the significant crop yield estimation models that are being used in agriculture. These models include nonlinear ones such as APSIM20, DSSAT21, 22, RZWQM23, and SWAP/WOFOST24, which, based on physiological data and soil processes, predict crop yield and other phenotypes. These models provide explicit interactions between traits and environmental conditions in different phases of the crop growth cycle (Ansarifar *et al*., 2021). However, the collection of trait measurement data and the calibration of model coefficients can be a challenge due to the intensive labor and time required in developing them (Lamsal *et al*., 2017). As a result, the models have low computation speeds and low accuracy, unlike some machine learning algorithms.

Another commonly used method is to empirically relate remote sensing data to crop yields on a local or regional scale. These types of relationships are always investigated through the use of some indices generated from remotely sensed imagery. Wang *et al*. (2021) researched the assimilation of remote sensing data

into crop growth models in India. In this study, the potential of combining high-resolution LAI data with crop modeling to assess crop yields at field scale was investigated. The study confirmed that with the assimilation of remotely detected satellite data, crop yields could be easily predicted under rainfed and irrigated conditions using field data (Wang *et al*., 2021). Remote sensing data uses three types of crop yield estimation methods: (1) empirical models based on vegetation indices (VIs). García-Martínez *et al*. (2020) (2) crop yield estimation where mechanistic models are combined with remote sensing data (Wang *et al*., 2021); (3) semi-empirical production estimation models that use gross primary production (GPP) or net primary production (NPP) for crop yield estimation. The research indicates that the estimation precision of net primary productivity is significantly influenced by the type of models and input of key surface parameters of ecosystems (Wang *et al*., 2020). Furthermore, it indicates that advancements in remote sensing data and the rapid development of remote sensing data processing technologies have resulted in NPP estimation models based on remote sensing data. As such, there has been improvement in the NPP estimation models (Wang *et al*., 2020).

Studies have highlighted that plant physiology, stress, and yield capabilities are expressed in the spectral reflectance from crop canopies and could be quantified using spectral vegetation indices (Wang *et al*., 2021). Vegetation indices use empirical models' due to their simplicity in estimating crop yield estimation (Wang *et al*., 2021). This method generally uses the vegetation index at a single growth stage or the combined index to enable multiple growth stages to build univariate or multivariate models for yield estimation (Dente *et al*., 2008). Vegetation indices, notably the normalized difference vegetation index (NDVI), have been used in the estimated study. It is noted that "Chlorophyll "a" and "b" in the palisade layer of healthy green leaves absorb most of the incident red radiant flux while the spongy mesophyll leaf layer reflects much of the near-infrared radiant flux" (Shen *et al*., 2009). The NDVI reflects the relationship between healthy green vegetation and the spectral reflectance of near-infrared and red wavelengths to determine crop

green vegetation health and volume. This particular research study used multiple parameters in the estimation of sorghum and green gram yield. The generation of yield from satellite, soil, and relief data is crucial since each plant species depends on complex factors for yield formation and differs for every crop type (Lang *et al*., 2023). However, research on this topic has indicated that remote sensing alone is generally not capable of producing accurate yield estimations (Ruby 2002). This has led scientists to look for other techniques that can be combined with remote sensing data to give better results. Present crop yield estimation uses methods and data sources like field surveys, expert knowledge, trend analysis, regression analysis, statistical models, and crop growth simulation models (Hoogenboom *et al*., 2019). In the early 1980s, Tucker (1980) found that arithmetic calculations of vegetation reflectance in the red and near infrared are particularly useful for vegetation classification. The index became most popular for studying vegetation health and crop production. The success of the NDVI is a result of the canopy leaf area index (LAI) and fAPAR (fraction of absorbed photosynthetically active radiation) (Myneni and Williams, 1994). The linear relationship with fAPAR and the NDVI enables an indirect measure of primary productivity.

One obstacle that affects modeling and prediction of crop yields using remotely sensed data is the classification of image masks (Wardlow *et al*., 2007). Image masking, which involves analyzing a subset of pixels in a region during yield prediction, has been found to be important. For instance, Doraiswamy *et al*. (2003) used three years of AVHRR NDVI imagery to assess spring wheat yields in North and South Dakota in the US. They concluded that the most promising way to improve the use of AVHRR NDVI for estimating crop yields is to use better crop masks. Working with medium- or coarse-resolution images (about 25–100 ha/pixel) for the development of crop masks is difficult and not feasible at regional scale. This is particularly true in low-producing regions with a sparse crop distribution. The unmixing of medium- and coarse-resolution data for a specific crop is challenging, as demonstrated by Atzberger (2013) and Rembold *et al*. (2013). The combined use of

high-resolution and medium-resolution data for crop area estimation is further demonstrated by Zhang *et al*. (2013). A more feasible alternative is cropland masking, which refers to the use of pixels dominated by "arable land," as described by Doraiswamy *et al*. (2003). This approach neglects crop-specific growth patterns. To overcome the shortcomings related to cropland masking and crop-specific masking (Wardlow *et al*., 2007), they proposed a new masking technique called yield-correlation masking. Yield crop correlation masking is challenged by the change in the selected landcover for yield. This requires the use of crop models in remote sensing to improve the accuracy of information (Wardlow *et al*., 2007).

2.8 Crop Modeling

Modeling is a way to simplify a system, where system is defined as a part of reality that contains interacting components (Liang *et al*., 2008). Crops are populations of plants, of which the growth is managed by humans for any of the various uses they may have for (parts of) these plants (Jones *et al*., 2017). Crop modeling is a way of simplifying the part of reality (i.e., system) known as a crop (Hoogenboom *et al*., 2019). At first, crop models were developed mainly to increase understanding of the basic processes of crop growth and development (Ansarifar *et al*., 2021). The increasing number of applications resulted in the development of a multitude of different models, and soon the need was felt to combine different models in one single framework for handling the modeling and analysis needs of different cropping systems in different environments (Hoogenboom *et al*., 2019). Two examples of such simulation frameworks are Decision Support System for Agro technology Transfer (DSSAT)(Boote *et al*., 1996), and Agricultural Production Systems Simulator (APSIM), (Jones *et al*., 2017). They provide structures to easily incorporate new models and to enable the simulation of different crops. The diversity of models leads to a diversity of reuses and abuses of models (Boote *et al*., 1996). Models are used to enable the simulation, extrapolation, description and understanding the functions of a dynamic systems (Lamsal *et al*., 2017). This is not different for crop models: they can be and are, used for various objectives (from

understanding crop functioning to scenario analysis exploration) by different users (from researchers to policy makers) and at various scales (from gene to globe). Because of this plethora of models and their uses, the quest for balance in crop modeling between the objectives of the simulation and the approaches selected has become a subtle exercise (Hoogenboom *et al*., 2019).

Crop growth modelling started in the late 1960s (Makowski *et al*., 2006) . Initially, crop models were developed to increase understanding of the mechanisms (i.e., physiological processes) underlying crop growth and development. Subsequently, the focus in crop modelling was the understanding to which main abiotic factors (weather and soil) constrain crop yields (Boote *et al*., 1996). As a result, each crop growth model tended to focus on one crop and one specific set of conditions. Different modelling approaches have been developed to simulate the same process (Lamsal *et al*., 2017). For example, various algorithms have been used for the simulation of biomass accumulation. Some models comprise a very detailed description of the processes related to photosynthesis and respiration Jones *et al*. (2017) while others use the radiation use efficiency approach Hoogenboom *et al*. (2019) representing the detailed photosynthesis and respiration models by one parameter. Similar differences in modelling approaches and detail can be found for other crop physiological processes(Lamsal *et al*., 2017). Therefore, there is a need for balanced crop modelling where the focus is the understanding of the impact of a wide range of conditions, while at the same time employing empirical approaches with the purpose of estimation, but specific to a particular set of (local) biophysical conditions (Husak *et al*., 2008).

Empirically derived agricultural models contain many simple relationships that do not completely represent actual plant processes. Testing of models against observed data (calibration and validation process) provides data that can be used to understand crop situations under current climate conditions (Makowski *et al*., 2006). However, problems can arise from the simplification of the crop models. For example, agricultural models make the assumption that weeds, diseases, and insect pests do

not exit or are controlled. Jones *et al*. (2017), that the soil conditions are adequate for the crop grown (Brisson *et al*., 2003), and that the weather events are within the normal range (Adam *et al*., 2012). A range of agricultural models are widely used by scientists, technical extension services, commercial farmers, and resource managers (Adam *et al*., 2012).

CHAPTER THREE

MATERIALS AND METHODS

3.1 Description of the study site

This chapter outlines a description of the study area, research designs used to achieve each research objective, data collection methods for secondary and primary data, and methods used in statistical data analysis. The research was carried out in Chiakariga, Tharaka Nithi County, and Katangi-Ikombe Wards, in Machakos County of Kenya, for the year 2015,2017,2018,2020 and 2021 as shown in Figure 3.1. This research covered both sorghum and green gram.

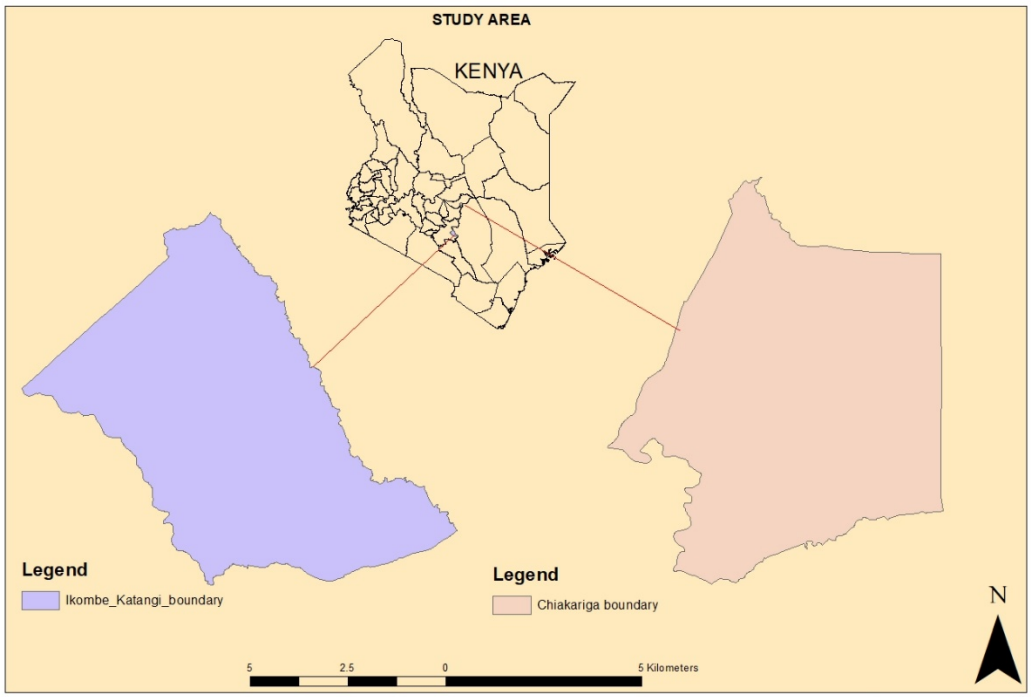


Figure 3.1: Map of the Ikombe-Katangi Ward and Chiakariga study area

3.2 Estimation of spatial distribution and spectral signatures of sorghum and green gram.

3.2.1 Spatial data for sorghum and green gram

The study presents a multifaceted approach to identifying green gram-growing areas in semi-arid areas of agroecological zones IV and V (Figure 3.1) using a suite of methods, including reflectance data and image statistics, and temporal-based approaches to the identification of cropping patterns and spectral signature ranges for crop identification in specific crop stages, seasons, and agroecological zones. The study used a time series of 16-day, 30-m spatial resolution, seven-band reflectance data from the medium-resolution imaging spectroradiometer (Landsat OLI8) sensor, as well as ancillary spatial data sets and field plot observations, to identify and classify green gram areas over a large spatial extent. In this study, the demarcation of cropped areas was applied through field survey data and the establishment of crop spectral signature ranges for the same field in a particular season for a particular phenological stage to enable the phenological cycle and their statistical (di)similarity. Furthermore, the landscape vector data for the study area was acquired from the Kenyan administrative ward data to help extract the area of interest in the satellite imagery. The study used data from 2013, 2015, 2017, 2018, and 2020 to build the historical data required for spectral reflectance calculation while ensuring the selection of days with minimal to no cloud cover for the study area during the OND rains period where green gram and sorghum cultivation was carried out. To counter the challenges of non-uniformity of coverage and seasonal differentiation of crops, the research acquired images with similar temporal resolution, working in a specific agroecological zone for one phenological stage of crop in a particular growing season. The following equipment was used in the field:

1. GPS device-Trimble Juno
2. Satellite imagery (medium - Landsat 8)
3. GIS, Remote sensing and Statistics Software (R, QGIS, ARCGIS, ENVI and Terrset)
4. Soil Auger

3.2.2 Remotely sensed image data

The research used Landsat 8 data for the October November December rain season for 5 years i.e., 2020,2018,2017,2015 and 2013 within the study area. Multi-temporal Landsat data have been found to be very useful in crop monitoring. The NASA satellites Landsat-8 provide images every 16 days with a spatial resolution of 30 m. Acquisition of data considered images within the same temporal resolution.

3.2.3 Field data collection

The field data collection followed on farm correlational research design, (Kitrell,1974) at and farmers surveys. The following activities were carried out in the field to enable the field research study to complement the data collected through remote sensing. The GPS coordinates extracted from the field survey from the year 2020 was instrumental in generating data that enabled cropping pattern data assignment in ARCGIS software on the various parcels of farm.

i. Farm parcel digitization
ii. Setting up green gram farms for data collection (long and short season)
iii. Collection of historical data on green gram farms
iv. Soil Data Collection and Analysis
v. Acquisition of satellite imagery data Landsat 8(2013,2015,2017,2018 and 2020)

3.2.4 Training data

Ground observation data for the various cropping patterns was collected for the year 2020, and a socioeconomic survey (Appendix 1.0 questionnaire) was conducted to collect more data for the cropping patterns undertaken for the last 10 years. These sets of data were important in understanding the changing cropping patterns across small-holder farms.

3.2.5 Soil Data Collection and analysis

Soil data was collected for all 15 farms that formed the controls within the study areas. The soil collected was analyzed for soil pH and soil nitrogen, potassium, and phosphorus at KARLO Mwea, and Figure 3-3 shows the soil sampling data collection.

A **B**

Figure 3.2: Soil sampling in Chiakariga(A) and Ikombe-Katangi(B) study areas

3.2.6 Setting up farm control in the field.

The method used for data collection and setting up the farm controls was the multistage random sampling design, where a homogeneous population was generated through the elimination of farms that were not representative of the study (Figure 3.2 below). The experiments was done according to Stafford *et al*., (2006) who used multistage sampling on rice fields to select plots in the landscape that had a population of rice farmers. This method used all digitized farms within the study area for the sampling design (Figures 3.6 and 3–7). The farms were digitized from online gazetteers (Google Earth). The number of farms in the study area was 10,000, which is equivalent to the total number of farms confirmed from the field survey. Figures

3.2. Level 1 of the sampling involved the elimination of farms under irrigation to remain with rainfed farms only. The irrigated farms were 5% of the total farms within the study area. The irrigated farms were identified through the Normalized Difference Moisture Index, which is calculated using the Near Infrared (NIR) and Short-Wave Infrared (SWIR) as follows:

NDMI = (NIR – SWIR) / (NIR + SWIR) …….3-1

Normalized Difference Moisture Index (NDMI) is used to detect moisture level in vegetation and therefore applied on the digitized farms in the dry season to identify farms under irrigation. This enabled the identification of irrigated farms using sentinel 2 satellite and these farms were eliminated from the sample. Decision tree algorithm in ARCGIS software was used to preselect the rainfed farms only.

This resulted in n = 9,500 farms available for sampling. Level 2 of the sampling involved the elimination of farms that did not have any crop planted or where the crop was planted late within the planting season in question, i.e., short rains (OND) from the total population of farms under level one. Investigation into farms that were not cropped was carried out using Sentinel 2 satellite data, where the Normalized Difference Vegetation Index (NDVI) was run for the study area. The NDVI was used to detect vegetation on the digitized farms in the study area. This resulted in 35% of 9500 farms from level one being eliminated, which resulted in 6175 farms. At level four of the sampling, sublocation administrative data was used to select areas where green grams had been grown in the last five years. The first level involved interviewing the government extension officer to understand administrative areas where green gram is grown within the study area. Six sublocations out of the total of 10 were selected, resulting in 3705 farms. Of the 3705 farms, another criterion selected farms where green gram had been grown for at least ten consecutive years. Identification of farms where green gram and sorghum have been grown was done through conducting a survey of three questions (whether the farmer has grown green gram or sorghum, the frequency of growing green gram or sorghum, and what season the crop was grown). The sample questionnaire can be found in Annex 1. This resulted

in 820 farms, of which 10% were sampled for data collection. The 10% sample selection was done using the formula.

These steps were also followed for the Chiakariga study site, resulting in the initial selection of 1014 farms. There were no irrigated farms or multiple sublocations growing sorghum in the Chiakariga area, but all other elimination criteria were followed. Of the total population, 875 farms were planted late or did not plant the sorghum crop. There were 771 farms left for sampling. Furthermore, farms that had grown sorghum for the last 10 years were selected. There were 123 farms that had grown sorghum consecutively in the last 10 years. This means that 648 farms were eliminated from the initial sample population. Of the 123 farms, 10% were sampled for data collection. 10% of the statistics are arrived at based on a sampling study by Cohen *et al*. (2007), who discuss in detail various research methods. In his study, he highlights that for a sample population less than 1000, a 10% population sample size is representative. The 10% representative sample was selected using the formula;

$n = [(N)\ (p)\ (1 - p)] / [(N - 1)\ (B/C)^2 + (p)\ (1 - p)]$…….3-2

Where n is the computed sample size needed for the desired level of precision; N is the population size; p is the proportion of the population expected to choose; B is the acceptable amount of sampling error, or precision; and finally, C is the Z statistic associated with the confidence level, which is 1.96 and corresponds to the 95% level. B was set at 0.05, which is ± 10, 5% of the true population value, respectively. The acceptable amount of sampling error or precision was set at 0.05 or 5%. A confidence level of 1.96 corresponding to the 95% level.

For green gram

Where N =820, p = 0.1, B = 0.05, C = 1.96

n = [(820) (0.1) (1-0.5)] / [(820-1) (0.05/1.96)2 + (0.5) (1-0.5)]

n= [(820) (0.1) (0.5)]/ [(819) (0.0255)2 + (0.5) (0.5)]

n= (41)/ (1.56+ 0.25)

n= 41/ 1.81

n= 22.65 ≈ 23 for green gram

For Sorghum

Where N =123, p = 0.1, B = 0.05, C = 1.96

n = [(123) (0.1) (1-0.5)] / [(820-1) (0.05/1.96)2 + (0.5) (1-0.5)]

n= [(123) (0.1) (0.5)]/ [(819) (0.0255)2 + (0.5) (0.5)]

n= (6.15)/ (1.56+ 0.25)

n= 6.15/ 1.81

n= 3.397 ≈ 4

The sampled farms for cropping pattern training data were 82 farms for green grams and 12 farms for sorghum (Figures 3.3 and 3.4, respectively). The sample size was increased to ensure enough ground truth data for satellite imagery training. Of the farms selected, there were 12 and 3 control farms for green gram and sorghum, respectively. A control farm involved having a pure stand of the crop in the field where all the good agricultural practices, such as the use of appropriate inputs (certified seed and fertilizers), weeding, and pest control.

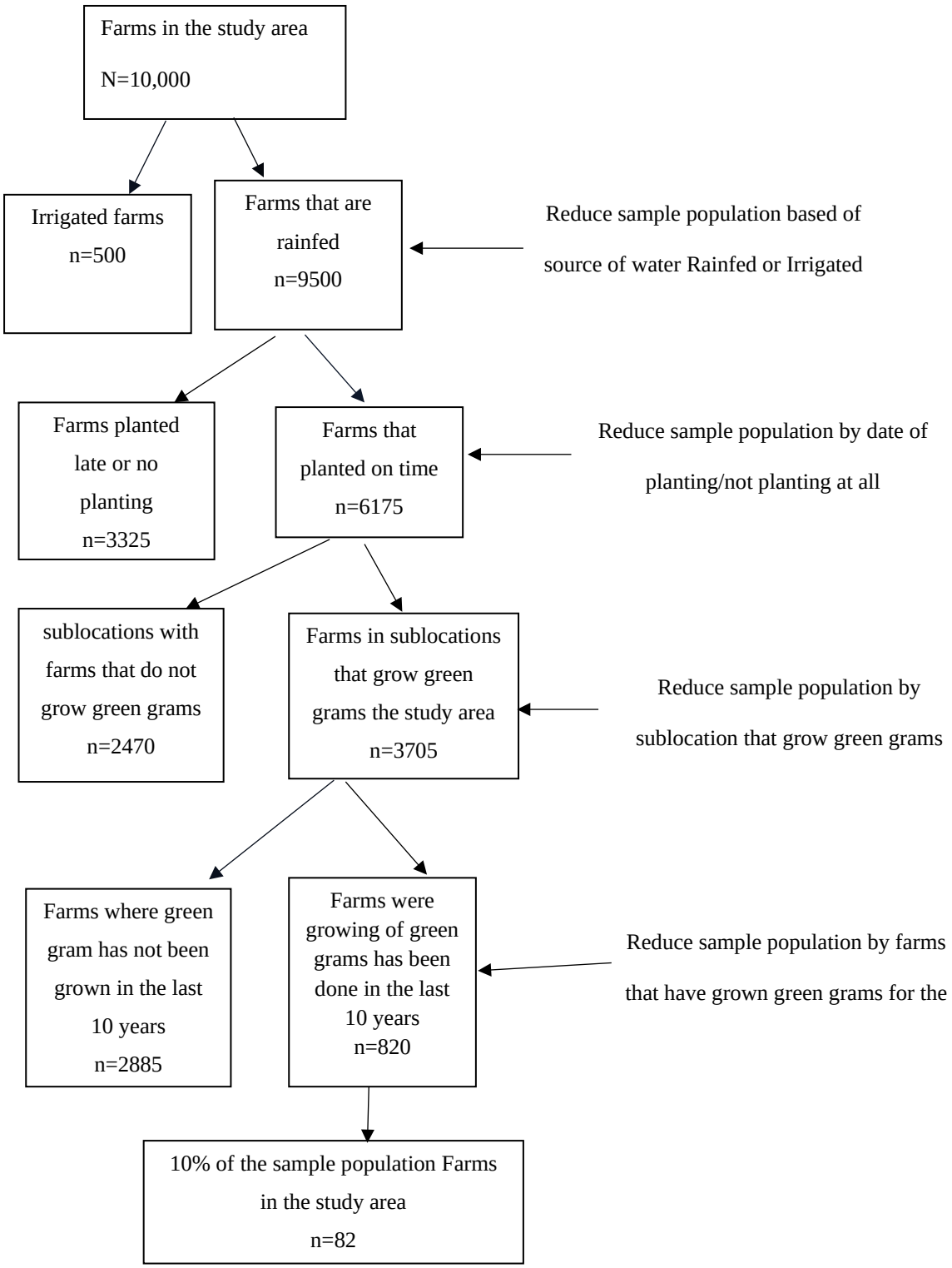


Figure 3.3: Random sampling design for the Ikombe-Katangi study area

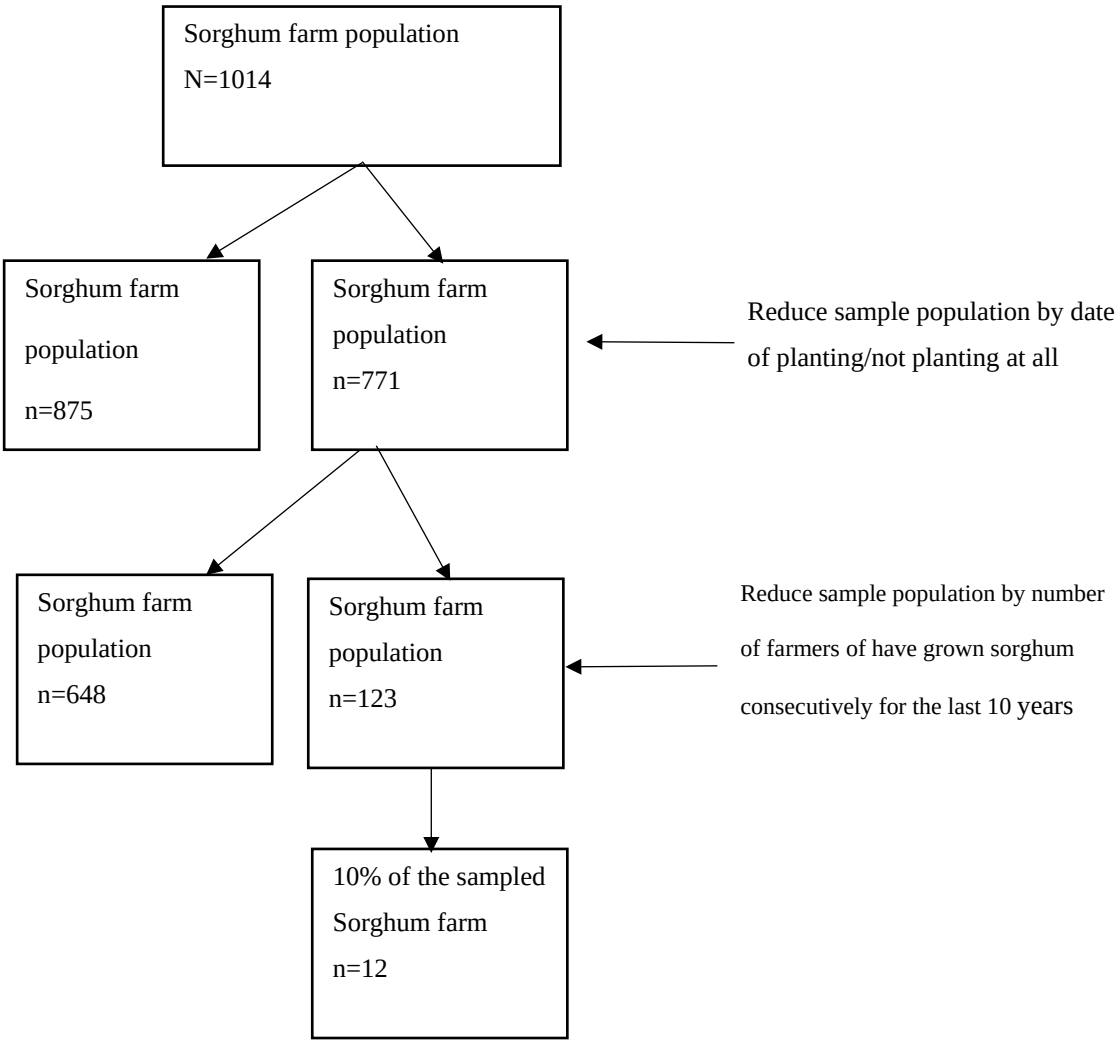


Figure 3.4: Random sampling design for the Chiakariga study area

The field survey was set up for the 2020 (OND) rains and the 2021 short rains. This exercise involved planting certified seeds, fertilizer application, weeding, and generally monitoring this crop from planting to harvesting (Figure 3.4 for field crop monitoring photos).

Figure 3.5: Monitoring of the crop of sorghum (A) and green gram (B) in the field

3.2.7 Spatial distribution and crop spectral signature assessment for sorghum and green gram

3.2.7.1 Digitization of farm parcel boundaries using Google Earth engine

Farm parcels within the study area were identified and GPS coordinates were collected, and these were used to digitize the farm using an online gazetteer, i.e., Google Earth. Figures 3.6 and 3.7 illustrate the farm parcel digitized, which was used in the calculation of the estimated area for green gram and sorghum. The study therefore allowed the generation of spatial distribution data for green grams and sorghum for the five years investigated, i.e., 2015, 2017, 2018, 2020, and 2021, over large areas of agroecological zones IV and V, through the use of decision tree

algorithm classification and field-generated socioeconomic data depicting the growing patterns of farmers using multitemporal remote sensing data.

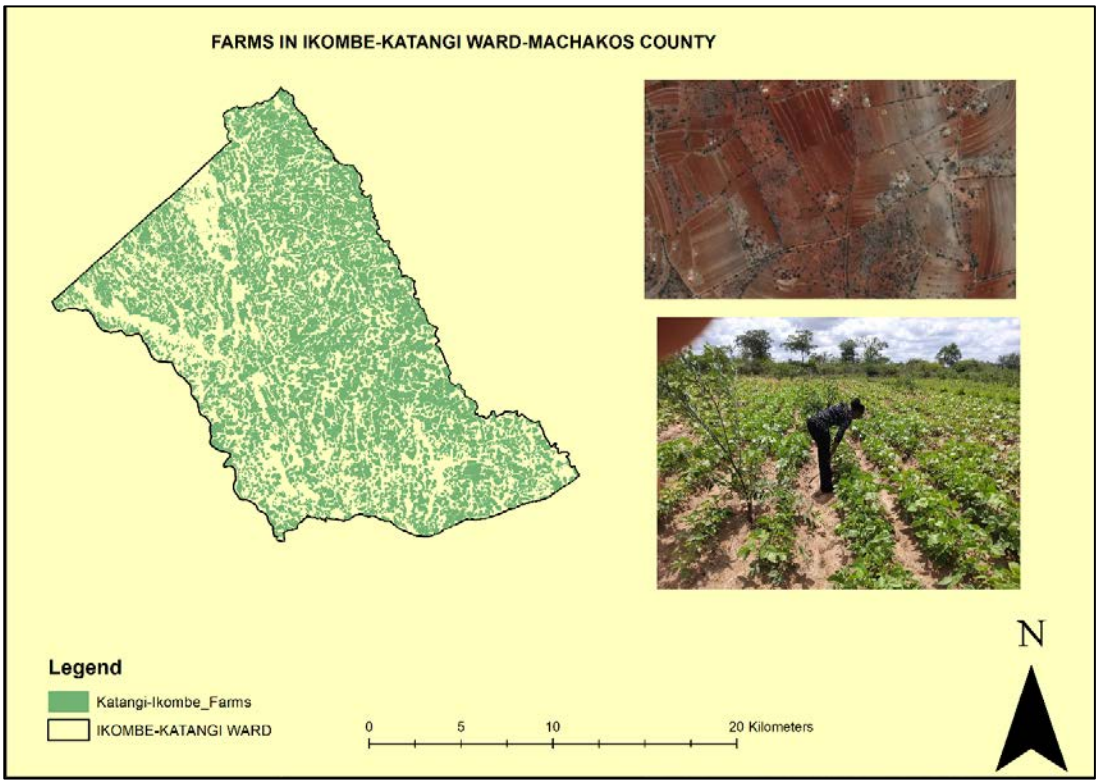


Figure 3.6: Digitized farm for Ikombe-Katangi, Green gram study area

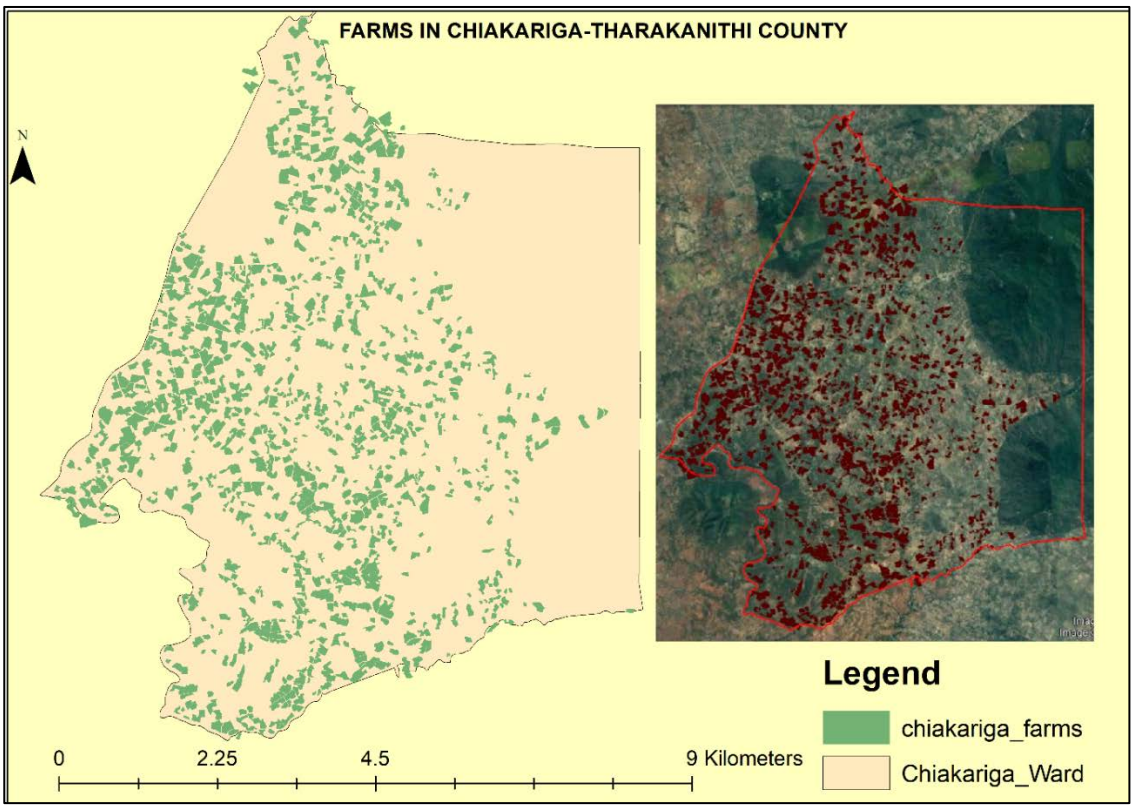


Figure 3.7: Digitized farms for Chiakariga Ward, Sorghum study area

This was a method of digitizing advanced area frame sampling by using a geographical information system in 4 stages. Stage one was automated object segmentation using high-resolution imagery such as online gazetteers like Google Earth to generate data that was used for farmed land sampling units based on the administrative units. The first step of the process was to use the feature statistics layers (i.e., NDVI, NDMI, and NIR) and threshold them into interest areas that approximated the boundaries of the thematic classes that were to be classified. This was done using fully convolutional networks (FCN) using MATLAB and R (Mudrik, 2019). Furthermore, this data was used as training data in the classification of the Landsat 8 OLI images to obtain farmed land within the administrative units. Once the farmed land was extracted and classified, the decision tree algorithms were applied to separate rainfed and irrigated farmed lands from the initial farmed land sampling units. Once the rainfed farms were removed, training data from the field survey along with algorithms by Mudrik (2019) was applied to identify sorghum and green gram farms

in the Landsat 8 OLI satellite image. This provided much-needed cropped data for sorghum and green gram within the study area. Spatial distribution was determined using supervised classification using a maximum likelihood classifier. The research used Landsat 8 data in all four stages of the assessment of the spatial distribution of green gram and sorghum.

3.2.7.2 Extraction of Green gram and Sorghum spectral signatures

This involved the calculation of the top-of-atmosphere reflectance from the original bands for the satellite images that were downloaded for 2015, 2017, 2018, 2020, and 2021. Table 3.1 shows the source, date, and time of the image. The calculation of TOA reflectance for bands 2, 3, 4, 6, and 7 is very important in the identification of cropped areas on Landsat 8 satellite imagery.

Table 3.1 Landsat 8 OLI scene satellite imagery metadata.

Landsat 8 product	Type of product	Path	Row	Date of acquisition	Cloud cover
LC08_L1TP	Collection 2 level 1	168	61	20/02/2020	1.30
LC08_L1TP	Collection 2 level 1	168	61	2018-01-29	0.02
LC08_L1TP	Collection 2 level 1	168	61	2017-12-28	0.26
LC08_L1TP	Collection 2 level 1	168	61	2015-01-05	2.48
LC08_L1TP	Collection 2 level 1	168	61	2013-11-15	3.59

A normalization procedure was used in order to remove variations between these images due to sensor differences, Earth-Sun distance, and solar zenith angle (Bruce and Hilbert, 2006). Conversion from DN to TOA reflectance was done using the equation illustrated below. Top-of-atmosphere reflectance was calculated by using the rescaling coefficients in the MTL format file found alongside the downloaded satellite images (USGS website), as shown in equations 3–3.

$\rho\lambda' = M\rho Qcal + A\rho \rho\lambda' = M\rho Qcal + A\rho$ ……3-1

where:

$\rho\lambda'$ = TOA planetary reflectance, without correction for solar angle. Note that $\rho\lambda'$ does not contain a correction for the sun angle.

$M\rho$=Band-specific multiplicative rescaling factor from the metadata (REFLECTANCE_MULT_BAND_x, where x is the band number).

Aρ =Band-specific additive rescaling factor from the metadata (REFLECTANCE_ADD_BAND_x, where x is the band number)

Qcal = Quantized and calibrated standard product pixel values (DN)

The TOA reflectance with a correction for the sun angle is then:

ρλ=ρλ′cos(θSZ) =ρλ′sin(θSE) ρλ=ρλ′cosf_0(θSZ) =ρλ′sinf_0(θSE)

where:

ρλ= TOA planetary reflectance

θSE = Local sun elevation angle. The scene center sun elevation angle in degrees is provided in the metadata (SUN_ELEVATION).

θSZ =Local solar zenith angle; θSZ = 90° - θSE

3.2.7.3 Spectral similarity assessment

All cropping pattern spectral signatures generated and identified from the ground observation were subjected to the Euclidean distance spectral similarity assessment by comparing all patterns against possible outcomes.

3.2.7.4 Supervised maximum likelihood classifier for land cover classification

In this study, a statistical conclusion rule that examines the likelihood function of a pixel for each of the classes and assigns the pixel to the class with the highest probability, a maximum likelihood classifier, was used (Richards, 1999). The classifier assumptions are based on training statistics; each class has a normal or 'Gaussian' distribution. The classifier then uses the training statistics to compute a probability value for whether it belongs to a particular land cover category class. This allows for within-class spectral variance. In this study, expert opinion knowledge was used to weight the probability function. The calculation of the variance-covariance matrix V for each class i in the Maximum Likelihood Classifier algorithm is illustrated in equation 1 below.

$$g_{i(X)} = \ln \rho(\omega_i) - {}^{1}/_{2} \ln|\Sigma_i| - {}^{1}/_{2} (x - m_i)T\Sigma_i^{-1}(x - m_i) \text{.....Equation 3-4}$$

$Where$:

$i = class$

$x = n - dimensional\ data\ (where\ n\ is\ the\ number\ of\ bands)$

$p(\omega i) =$

$probability\ that\ class\ \omega i\ occurs\ in\ the\ image\ and\ is\ assumed\ the\ same\ for\ all\ classes$

$|\Sigma i| = determinant\ of\ the\ covariance\ matrix\ of\ the\ data\ in\ class\ \omega i$

$\Sigma i - 1 = its\ inverse\ matrix$

$mi = mean\ vector$

3.2.7.5 Kappa coefficient statistics for accuracy assessment

The confusion matrix was performed by comparing the error values for each class that was classified with its respective value in the ground truth data. The ground truth points used for the accuracy assessment were obtained from the field during data collection. The kappa coefficient works in such a manner that the same number of columns and rows, which are equal to the number of classes, is achieved. The land cover classes in the ground-truth image head the rows, while the same classes for the classified image head the columns. The Kappa coefficient (K) is then calculated using the formula illustrated in the equation below:

$$K\wedge = \frac{N\sum_{1=1}^{r} Xii - \sum_{i=1}^{r} X_{i+X+i}}{N^2 - \sum_{i=1}^{r} X_i + X + i} \quad \text{.....Equation 3-5}$$

Where;

r is the number of rows/columns in the confusion matrix.

Xii is the number of observations in row i and column i

X_i+ is the total number of rows i

X+I is the total number of columns in column i

N is the number of observations.

3.3 Extraction of micro agroecological zones in the study area

3.3.1 Agroecological zones

This study focused on using the GIS-based multicriteria analysis (MCA) approach to define the subzones of agroecological zones IV and V in Kenya. This was achieved by applying the agroecological zonation criteria using remote sensing data. The process defined any changes that may affect the suitability of land for the growth of green gram and sorghum in AEZ IV and V. The research adopted additional parameters recommended by Manzi and Gweyi-Onyango (2020) to accommodate any changes that may result from climate change. The data used to establish the microclimatic zones in AEZ zones IV and V were soil, climatic, and topographic data. The soil data was composed of annual soil moisture averaged over 15 years, soil temperature averaged for 15 years, soil drainage, soil pH, and soil texture. The climatic data considered in the study were the precipitation average annually for 15 years and the land surface temperature annual average for 15 years. Finally, the topographic data was composed of slope and elevation.

3.3.1.1 Soil Characteristics

The data on soil characteristics was used in the study area, including soil pH, soil texture, soil moisture, soil temperature, and soil drainage. Data sets for soil characteristics were extracted from the International Soil Reference and Information Centre (ISRIC) database and the Goddard Earth Science Data and Information Services Center (GIOVANNI)/National Aeronautics and Space Administration (NASA) website for soil moisture and soil temperature.

3.3.1.2 Climatic parameters

In this study, the climatic parameters that were used were rainfall and land surface temperatures. Rainfall data was extracted from the Climate Hazards Group InfraRed

Precipitation with Station Data (CHIRPS) website for the study areas in both sorghum and green gram. Data processing was done in an ARCGIS environment. The land surface temperature was extracted from the GIOVANNI/NASA website for the study regions for both green gram and sorghum.

3.3.1.3 Topographical data

The topography of the study area refers to the slope and elevation properties of the land in the study area. Slope and elevation were calculated using Digital Elevation Models (DEM), which were downloaded from the USGS Earth Explorer in the ArcGIS environment. The Universal Transverse Mercator (UTM) projection and the WGS 84 datum (WGS 84 37S) were used in ArcGIS to make the topographical maps and fix the projection.The slope was determined by calculating the maximum rate of change between each cell and its neighbors. In the output raster, each cell had a slope value.

3.3.2 Mapping of micro-agroecological zones

3.3.2.1 Multi-criteria Decision-Making GIS Analysis Tool

This approach typically involves the use of multiple-criteria decision analysis (MCDA) models or decision rules for tackling spatial problems such as site selection problems and land use and suitability analyses (Ehrgott *et al*., 2010; Manzi and Gweyi-Onyango 2022). To successfully implement this assessment of subzones, an optimization model whose distinguishing characteristics have a geographical model was used. The quantity to be optimized is defined as the criterion in these types of models.

The subzone was undertaken by carrying out the assessment of land suitability in AEZ zones IV and V. This process involves known land suitability analysis, as documented by Khan *et al*. (2022). The FAO land suitability assessment considers various parameters, such as climatic, geographical, soil, vegetation, and other characteristics of the land, to determine suitable lands for specific uses. To evaluate the land

suitability of pulse crop land (green gram), eleven criteria were considered: topography (slope and elevation), climate (rainfall, land surface temperature), soil characteristics (soil texture, soil drainage, soil moisture, soil pH, soil depth, and soil temperature). Sorghum crop multicriteria analysis picked parameters similar to green gram crop but differed in the weighted scores assigned for the various parameters. The multicriterial overlay analysis methodology is illustrated in Figure 3.8.

The criteria used in the analysis were classified according to expert opinion and the literature. It was seen that a similar approach was used by Akinci *et al*. (2013), Everest *et al*. (2021), Mustafa *et al*. (2011), and Topuz and Deniz (2023), who used expert opinion in undertaking agricultural land suitability using GIS and the AHP process. In this study, the expert opinions were sourced from government agriculture extension officers who had experience and information having worked in the study area for more than 10 years. The weights of the criteria were assigned based on the experiences of agricultural extension officers within the study area, supported by different scientific literatures through pairwise comparison following the Analytical Hierarchy Process (AHP). AHP is a widely accepted decision-making method (Eskandari *et al*., 2012) that constructs a pairwise comparison matrix by assigning a value in the range of 1–9, as shown in Table 3.1 (Saatsaz *et al*., 2018). This method follows that, by assigning values for each factor against every other (Saatsaz *et al*., 2018), it finally gives an eigenvector weight, indicating that all criteria were ranked according to their significance following expert opinions and literature. Several studies have proven that the AHP process works in building criteria for land suitability assessment as well as addressing other land issues. The MCDA GIS analysis process for this study is summarized in Figures 3–6 below.

Table 3.2 Analytical Hierarchy Process (AHP) process

Intensity of Importance	**Definition**	**Explanation**
1	Equal importance	Two activities contribute equally to the objective.
3	Moderate importance of one over another	Experience and judgment slightly favor one activity over another
5	Essential or strong importance	Experience and judgment strongly favor one activity over another.
7	Demonstrated importance	An activity is strongly favored, and its dominance is demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation.
2,4,6,8	intermediate values between the two adjacent judgments	When compromise is needed
Reciprocal of above non Zero Number	If activity ‘m’ has one of the above non-zero numbers assigned compared to activity “n”, then n has the reciprocal value compared with m.	

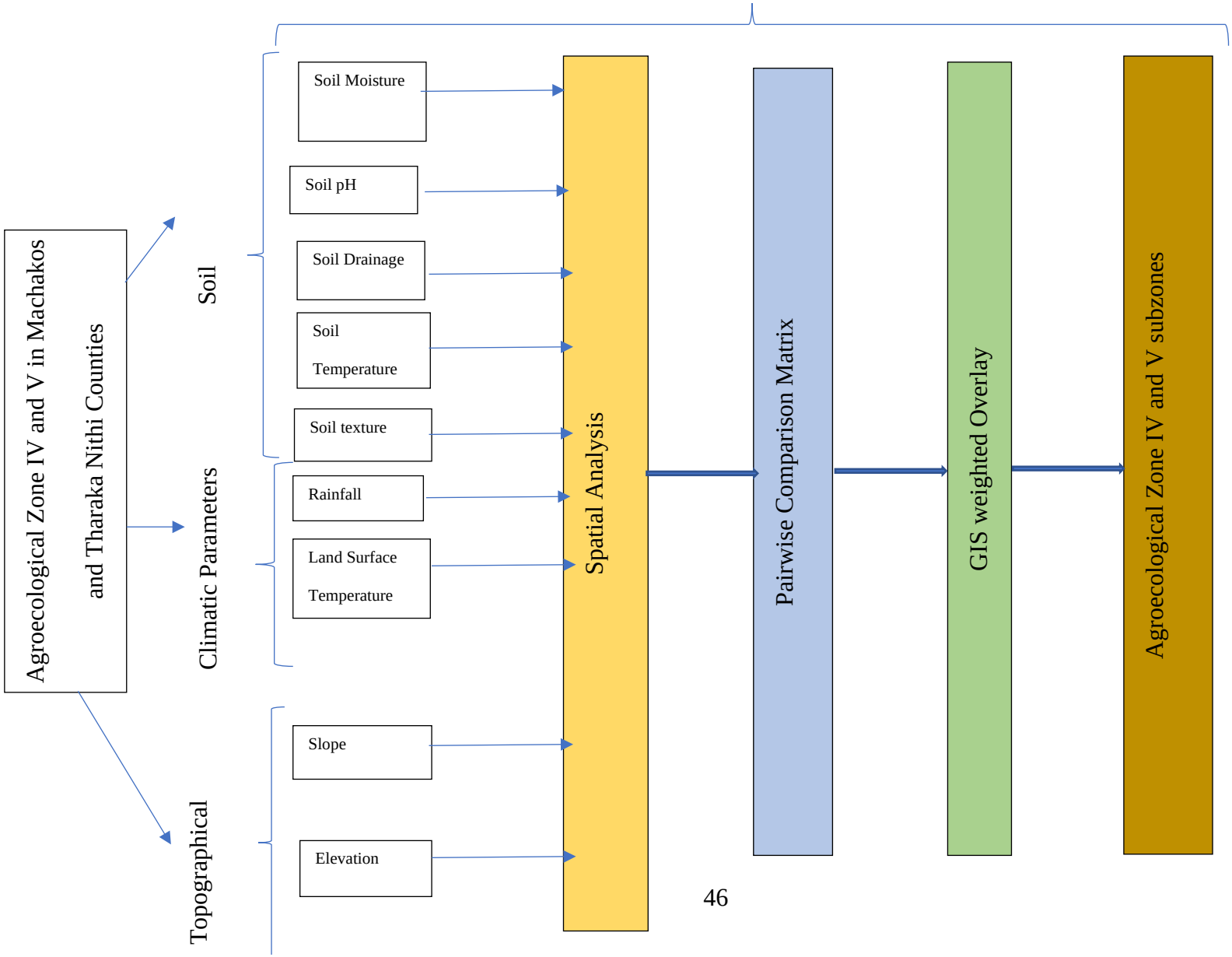


Figure 3.8: Subzone Multicriteria Overlay Analysis Methodology

3.3.2.2 Reclassification of parameters

The raster layers were reclassified in order to create a classification parameter that would enable the weighted overlay analysis for the agroecological subzone or microclimatic zones. Based on the AHP process of assigning eigen vector weights, FAO criteria on land suitable for agriculture, especially sorghum and green gram, were used for the reclassification of all parameters. This resulted in the following classes: very poor, poor, moderately poor, good, moderately good, and very good. These classes were used as a prerequisite for MCDA analysis to classify all parameters. Finally, a score was assigned to each class of each parameter. The reclassified criteria illustrated the area and spatial distribution of the various suitability levels of the criteria. Priority levels among the criteria were determined with the help of AHP analysis, and a weight was assigned to each criterion. This was followed by the deployment of a weight overlay in ArcGIS and the development of a final subzone map.

3.4 To estimate and model yields of sorghum and green grams in varied AEZ zones

3.4.1 Parameters used for crop yield estimation for green gram and sorghum crops

The study used parameters derived from remote sensing for the estimation of the green gram and sorghum yield. The research improved on crop yield estimation by taking into consideration crop yield models alongside other important datasets such as crop signatures, micro-agro-ecological zones, and prior crop mapping to specifically identify cropland areas. The following parameters were used in the calculation and estimation of the model:

i. Soil Moisture at 10- 40cm
ii. Soil type (from existing soil database)
iii. Precipitation amount
iv. Leaf area index

v. Enhanced Vegetation Index
vi. Land Surface Temperature
vii. Nutrient status (soil fertility) that is, measure of N, P and K
viii. Crop biomass
ix. Crop area determined from area estimation in the previous methodology
x. Crop type (determined from the second methodology that establishes the classification map of the crop varieties through crop signatures)

These parameters were evaluated in the vegetative stages of growth for each of the crops selected for the research, as mentioned in the previous methodology on crop variety discrimination. The selected phenological stage of growth was determined based on the critical stages of growth for each crop, that is, sorghum and green grams, which are likely to greatly influence crop yield. The assumption was that, in these micro-ecological zones, variability in crop management is at its minimum and homogenous fields can be achieved. The yield estimation involved the extraction of data for the various parameters mentioned above through the methods discussed for each parameter.

3.4.1.1 Soil Moisture

In this study, we used Landsat 8 and sentinel TOA data to retrieve the daily estimation of soil moisture patterns. Soil moisture was developed using the perpendicular soil moisture index (PSMI), which quantifies the status of soil moisture using reflectance values in the red, near-infrared, and thermal bands of satellite imagery. The ground cover and the Thermal Infrared Red (TIR) defined the trapezoidal space that represents the wet (db) and dry (c-e) edges of soil moisture.

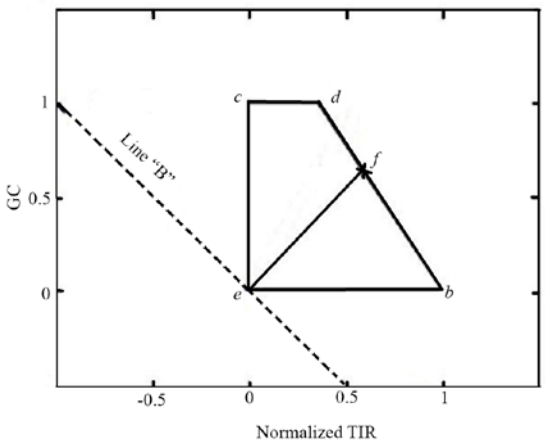


Figure 3.9: Perpendicular soil moisture graph. Source Shafian and Maas, (2015)

A straight perpendicular line e-f (Di) with a slope of -1 between the wet and dry edges was used against the ground cover to estimate soil moisture. Perpendicular Soil Moisture Index (PSMI) The thermal InfraRed (TIR) variation is mainly influenced by soil moisture, and the relationship between the index and soil moisture is due to fluctuations in the thermal emittance of the surface and the proportion of ground cover. The process involved the following steps: performing a cell statistic of the thermal infrared bands and consequently normalizing the resultant TIR raster.

$$TIR_{i,norm} = \frac{(TIR_i - TIR_{min})}{TIR_{max} - TIR_{min}} \ldots\ldots\ldots\ldots\ldots 3-6$$

The distance D*i* is

$$D_i = \left(TIR_{i,norm} + GC_i\right)/\sqrt{2} \ldots\ldots\ldots..3-7$$

The index:

$$PSMI_i = \frac{D_i}{1 + GC_i} \ldots\ldots\ldots\ldots\ldots.3-8$$

3.4.1.2 Soil Type

The study used the latest spatial soil database for Kenya, SOTWIS from ISRIC data hub (Batjes *et al*., 2007),This harmonized set of soil parameter estimates for Kenya (KENSOTER), at scale 1:1M, compiled by the Kenya Soil Survey. The parameters estimates are presented by unit of soil for fixed depth intervals of 0.2 m to 1 m depth for organic carbon, total nitrogen, PH (H2O), CEC soil, CEC clay, base saturation, effective CEC, aluminum saturation, CaCO3 content, gypsum content, exchangeable sodium percentage (ESP), electrical conductivity of saturated paste (ECe), bulk density, content of sand, silt and clay, content of coarse fragments, and available water capacity.

3.4.1.3 Precipitation amount

The processing of the downloaded rainfall products from CHIRPS involved, converting the files to raster in ArcMap using the multidimensional toolbox or extraction of binary files using a software such as PowerISO, evaluation for completeness, projection to an appropriate coordinate system, rescaling to a common resolution and reclassifying to a similar extent for easy comparison of data from different sources. In this research, precipitation data from the CHIRPS website was downloaded and processed using ARCGIS software for use in estimating green gram and sorghum yield.

3.4.1.4 Leaf area index

In this study, the leaf area index was calculated using an approach by Kang *et al*. (2016). The method adopted was the EVI2 vegetation index. EVI2 is less saturated at high LAI ranges and is less sensitive to the soil background and atmospheric noise (Kang *et al*., 2016). It is also the strongest for crop-specific relationships. The formula for calculating LAI using the EVI2 vegetation index is illustrated in Equations 3–16 below. (Long and Eitel, 2008; Name and Range, 2004)

$$LAI = (5.3\sqrt{EVI2} - 1.66)^{\frac{3}{2}} \dots\dots\dots\dots\dots\dots 3-9$$

(Kang *et al*., 2016) were able to establish that saturation is less common in relationships using EVI and EVI2 compared to NDVI, and in some cases, EVI and EVI2 are linearly related to LAI, indicating an ability to resolve LAI differences over a wider range of canopy conditions. In this case, the global LAI-VI relationships that were built by Kang *et al*. (2016) in Table 3.3 show the LAI-VI relationships for specific crops as obtained from the study. It is from this table that the coefficients for maize and soy beans were used for sorghum and green gram, respectively.

Table 3.3: Crop coefficient for various crops vegetation indices. Source Kang *et al*., (2016)

Crop Type	VI	SLR Model	Coefficient (Confidence Interval)		Prediction Model	RMSE (m^2/m^2)	MAE (m^2/m^2)	Quantiles of Absolute Residuals (m^2/m^2)				
			a	b				5%	25%	50%	75%	95%
Overall	EVI	$\sqrt{y} = a\frac{1}{\sqrt[3]{x}} + b$	3.51 (3.35,6.67)	−1.22	$y = \left(a\frac{1}{\sqrt[3]{x}} + b\right)^2$	1.46	1.09	0.07	0.34	0.84	1.6	2.93
	EVI2	$\sqrt{y} = a\frac{1}{\sqrt[3]{x}} + b$	3.73 (3.56,3.89)	−1.32	$y = \left(a\frac{1}{\sqrt[3]{x}} + b\right)^2$	1.45	1.08	0.07	0.33	0.8	1.56	2.98
Row crop	EVI	$\sqrt[3]{y} = a\sqrt[4]{x} + b$	2.81 (2.68,2.95)	−1.02	$y = (a\sqrt[4]{x} + b)^3$	1.52	1.13	0.07	0.32	0.84	1.67	3.24
	EVI2	$\sqrt[3]{y} = a\sqrt[4]{x} + b$	2.97 (2.83,3.11)	−1.11	$y = (a\sqrt[4]{x} + b)^3$	1.52	1.12	0.06	0.32	0.8	1.65	3.32
Maize	EVI	$\sqrt{y} = a\sqrt{x} + b$	3.8 (3.57,4.03)	−1.03	$y = (a\sqrt{x} + b)^2$	1.37	1.01	0.04	0.31	0.74	1.43	2.87
	EVI2	$\sqrt{y} = a\sqrt[3]{x} + b$	5.06 (4.79,5.34)	−2.23	$y = (a\sqrt[3]{x} + b)^2$	1.31	0.94	0.03	0.26	0.66	1.31	2.85
Soybean	EVI	$\sqrt[4]{y} = a\frac{1}{\sqrt{x}} + b$	-0.51 (−0.57,−0.44)	1.88	$y = \left(a\frac{1}{\sqrt{x}} + b\right)^4$	1.07	0.82	0.04	0.24	0.68	1.3	2.04
	EVI2	$\sqrt[3]{y} = a\frac{1}{\sqrt{x}} + b$	−0.69 (−0.74,−0.56)	2.19	$y = \left(a\frac{1}{\sqrt{x}} + b\right)^3$	1.06	0.82	0.05	0.23	0.67	1.31	2
Wheat	EVI	$y^{\frac{3}{5}} = ax + b$	3.73 (3.37,4.07)	0.23	$y = (ax + b)^{\frac{5}{3}}$	1.3	1.04	0.89	0.4	0.88	1.55	2.52
	EVI2	$y^{\frac{3}{5}} = ax^{\frac{3}{5}} + b$	4.82 (4.36,5.28)	−0.87	$y = \left(ax^{\frac{3}{5}} + b\right)^{\frac{5}{3}}$	1.32	1.04	0.1	0.34	0.88	1.46	2.66
Rice	EVI	$\sqrt{y} = ax^{\frac{2}{5}} + b$	3.83 (3.06,4.46)	−1.05	$y = \left(ax^{\frac{2}{5}} + b\right)^2$	1.41	0.97	0.01	1.18	0.65	1.27	3.08
	EVI2	$\sqrt{y} = ax^{\frac{2}{5}} + b$	3.9 (3.14,4.57)	−1.04	$y = \left(ax^{\frac{2}{5}} + b\right)^2$	1.4	0.96	0.03	0.2	0.58	1.24	3.02
Cotton	EVI	$\sqrt[3]{y} = a\frac{1}{\sqrt[3]{x}} + b$	−1.28 (−1.4,−1.16)	3.03	$y = \left(a\frac{1}{\sqrt[3]{x}} + b\right)^3$	0.94	0.73	0.04	0.22	0.66	1.09	1.91
	EVI2	$\sqrt[3]{y} = a\frac{1}{\sqrt[3]{x}} + b$	−1.24 (−1.35,−1.11)	3.02	$y = \left(a\frac{1}{\sqrt[3]{x}} + b\right)^3$	0.97	0.77	0.03	0.28	0.66	1.15	1.99
Pasture	EVI	$\sqrt{y} = ax^{\frac{6}{5}} + b$	1.6 (1.4,1.81)	0.71	$y = \left(ax^{\frac{6}{5}} + b\right)^2$	1.05	0.84	0.08	0.38	0.7	1.12	2.1
	EVI2	$\sqrt{y} = ax^{\frac{4}{3}} + b$	1.79 (1.57,2.03)	0.75	$y = \left(ax^{\frac{4}{3}} + b\right)^2$	1.05	0.84	0.07	0.39	0.69	1.11	2.07

3.4.1.5 Enhanced Vegetation Index

The study acquired the images and corrected for reflectance using equation 3-10 and 3-11 below. Landsat 8 OLI band data was converted to TOA planetary reflectance using reflectance rescaling coefficients provided in the product meta-data file (MTL

file). The following equation was used to the convert the DN values to TOA reflectance for Landsat 8OLI data as follows:

pA' =Mp*Qcal + Ap …….3-10

where:

pA' = TOA planetary reflectance, without correction for solar angle. Note that pA' does not contain a correction for the sun angle.

Mp = Band-specific multiplicative rescaling factor from the metadata (REFLECTANCE_MULT_BAND_x, where x is the band number).

Ap = Band-specific additive rescaling factor from the metadata (REFLECTANCE_ADD_BAND_x, where x is the band number)

Qcal = Quantized and ca liberated standard product pixel values (DN)

TOA reflectance with a correction for the sun angle is then:

pA = pA'/cos(theta_sz) = pA'/sin(theta_se)……3-11

where

pA = TOA planetary reflectance

theta_se= Local sun elevation angle. The scene center sun elevation angle in degrees is provided in the meta-data (SUN_E LEVATION).

theta_sz = Local so lar zenith angle; theta_sz=90- theta_se

The atmospherically corrected image was cropped to the extents of the study area.

The EVI formula from equation 3-19 was applied to the georeferenced image to obtain EVI values of the image using any GIS platform ARCGIS10.8 or QGIS, etc.

EVI = 2.5 * ((Band 5 – Band 4) / (Band 5 + 6 * Band 4 – 7.5 * Band 2 + 1))….3-12

In ARCGIS 10.8 the EVI value was calculated using the model builder tool through the calculate value tool, which allows the use of the Python math module to perform more complex mathematical operations. Once the EVI values were obtained. The EVI formulas used were;

$EVI = G * \rho NIR - \rho RED\rho NIR + (C1X\rho RED - C2X\rho BLUE) + L$…….3-13

and can be rewritten as;

$$EVI = G * \frac{\rho NIR}{\rho RED} - \frac{1\rho NIR}{\rho RED} + \left(C1 - C2X\frac{\rho RED}{\rho RED}\right) + \frac{L}{\rho RED}$$

Where:

L is a soil adjustment factor,

C1 and C2 are coefficients used to correct aerosol scattering in the red band using the blue band

The ρ_{blue}, ρ_{red}, and ρ_{nir} represent reflectance in the blue (0.45-0.52μm), red (0.6-0.7μm) Near-Infrared (NIR) wavelengths (0.7-1.1μm)

G is a gain factor.

In Landsat 8 the equation would be formulated as follows;

$$EVI = 2.5 * (\frac{BAND5 - BAND4}{BAND5 + 6 * BAND4 - 7.5 * BAND2 + 1})$$

where NIR corresponds to the near-infrared band (LANDSAT band 4), RED corresponds to

The red band (LANDSAT band 3), BLUE corresponds to the blue band (LANDSAT band 1)

3.4.1.6 Land surface temperature (LST)

The method used was a single channel algorithm that uses one band in the thermal radiance region. For example, Landsat 8 has two thermal bands, band 10 and 11. The formula used to calculate the land surface temperature is as shown:

$$\boldsymbol{LST} = \frac{\boldsymbol{T_{ToA}}}{\{\mathbf{1}+\left[(\frac{\boldsymbol{\lambda T_{ToA}}}{\boldsymbol{\rho}})\boldsymbol{LnLSE}\right]\}} \quad \text{……..} \ 3\text{-}14$$

Where ρ **=h**$\frac{c}{\sigma}$ **…….3-15**

Where;

LST= Land Surface Temperature

TToA =Brightness Temperature of the top of the Atmosphere Radiance

λ= Effective Band Wavelength

LSE= Land Surface Emissivity

ρ= Power emitted from the surface of the earth

h = Planck's constant

c = velocity of light

σ= Boltzmann constant

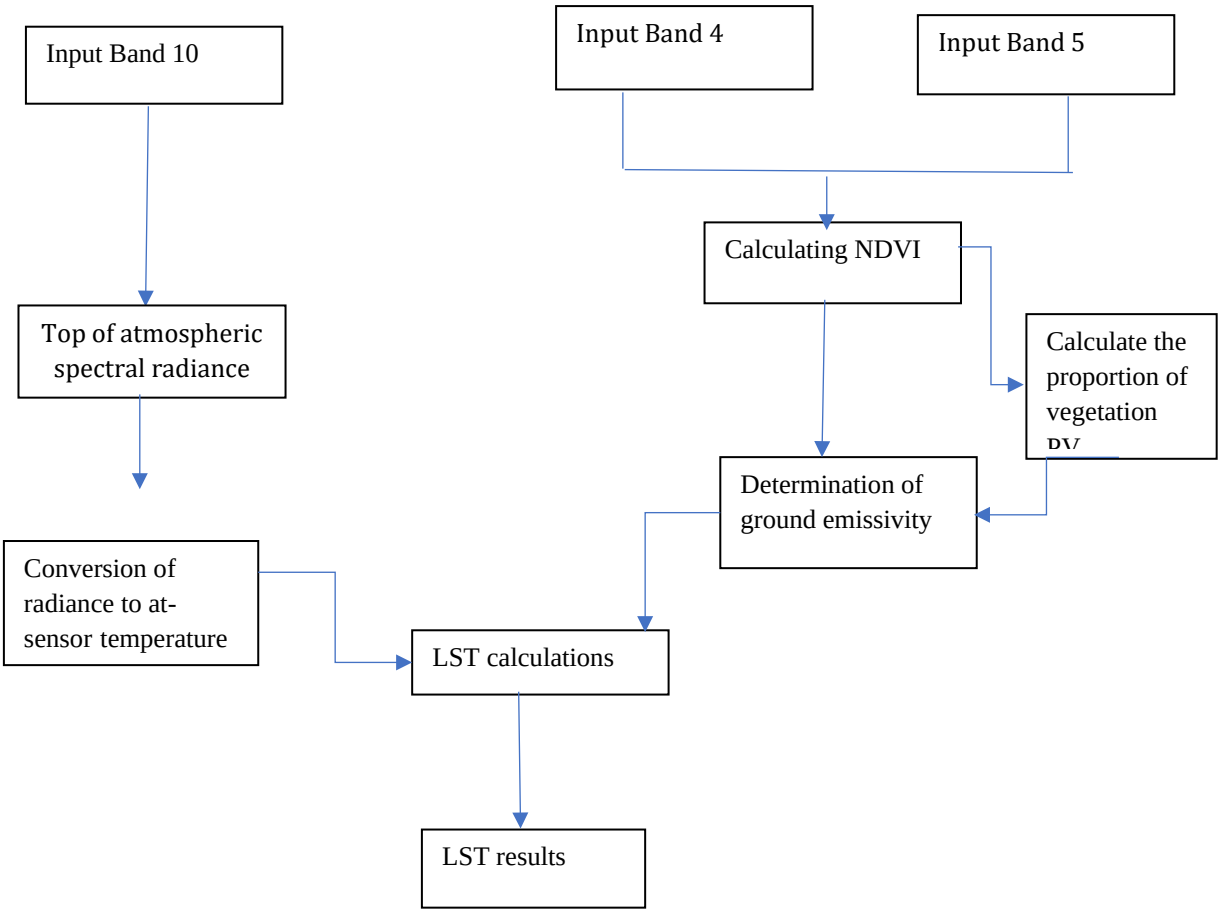


Figure 3.10: Summary of the process of obtaining land surface temperature from satellite imagery

The method used in the calculation of the temperature of the soil surface can be summarized in the following steps.

a. Normalized Difference Vegetation Index and the Proportion of Vegetation Cover

NDVI was used to express the amount of vegetation in a given area. It was calculated from the visible red and near-infrared light reflected by vegetation as shown in the following equation.

$$NDVI=(NIR-R)/NIR+R \quad \ldots\ldots..3\text{-}16$$

Using NDVI values, the proportion of vegetation cover was calculated as

$$\boldsymbol{P_y} = (\frac{\boldsymbol{NDVI_{NDVI_{min}}}}{\boldsymbol{NDVI_{max} - NDVI_{min}}})^{\boldsymbol{2}}$$……3-17.

where NDVImax = 0.5 indicates the presence of vegetation on the lands and NDVImin = 0.2 represents only bare soil on the surfaces of the land. NDVI value less than 0 indicates water, and NDVI value greater than 0.5 indicates full vegetation. NDVI values range between -1 and +1, which can include a mixture of vegetation. etc. This mixture of land surface characteristics is required for the calculation of P_v using the equation above.

b. Determination of Land Surface Emissivity

Emissivity is the ratio of the thermal radiation from a surface to the radiation from an ideal black surface at the same temperature as given by the Stefan-Boltzmann law (Wan and Zhao, 1997). The emissivity was calculated from reflectivity values in the red region of the image. where NDVI lies between 0.2 and 0.5, which was the case for the study area. The surface emissivity (LSE) for a given band i can be related to the NDVI and proportion of vegetation (Pv) (Vanhellemont, 2020), as shown in equation 4 below.

$$LSE_i = \begin{cases} a_i\rho_{red} + b_i & NDVI < 0.2 \\ \varepsilon_{v,i}P_v + \ \varepsilon_{s,i}(1 - P_v) + C_i & 0.2 \leq NDVI \leq 0.5 \\ \varepsilon_{v,i} + C_i & NDVI > 0.5, \end{cases}$$ ……3-18

Where;

εv,i- emissivity of fully vegetated surfaces

εs,i - emissivity of barren soil in band i

Pv - proportion of vegetation

ai, bi, the coefficients that can be estimated from laboratory spectra of soils using statistical fits, assuming that the emissivity and the reflectivity in the red band have a linear relationship.

Ci, the roughness of the land surfaces. For rough and heterogeneous surfaces such as soil-vegetation mixed pixels; Ci denotes the increment in emissivity resulting from the cavity effect and multiple scattering in the mixed pixels.

Taking emissivity values of soil and vegetation into account from the NDVI, the emissivity of land surfaces (LSE) were calculated according to the NDVI-threshold method as;

$$LSE = \varepsilon v\ Pv + \varepsilon s(1-Pv) + C \quad \ldots\ldots 3\text{-}19$$

The emissivities (εvand εs) were estimated using the MODIS UCSB (University of California, Santa Barbara, CA, USA) emissivity library. For example, the emissivity values for soil and vegetation for TIR bands 10 and 11 in Landsat 8 used are shown in the table 3.4 below.

Table 3.4: Emissivity values of soil and vegetation

TIR band	**Emissivity Values**	
	Vegetation(εv)	**Soil(εs)**
Band 10	0.9863	0.9668
Band 11	0.9896	0.9747

From the above approximation were done using the following equation.

$C = (1- \varepsilon s)\ (1-Pv)\ F\ \varepsilon v \quad \ldots\ldots\ldots 3\text{-}20$

where F is a shape factor that was considered by under different geometrical distributions having a mean value of 0.55(Sobrino and Jiménez-Muñoz, 2005).

Taking both equations into account,

$LSE = mPv+n \quad \ldots\ldots 3\text{-}21$

Where;

$m= \varepsilon v- \varepsilon s-(1- \varepsilon s)\ F\ \varepsilon v$, and $n= \varepsilon s+(1- \varepsilon s)\ F\ \varepsilon v'$

c. Determination of Top-of-Atmosphere (ToA) Brightness Temperature

The Top of Atmosphere (TOA) Reflectance brightness temperature is a unitless measurement which provides the ratio of radiation reflected to the incident solar radiation on a given surface**.** The brightness temperature is uniquely related to the radiance for a given wavelength by the Planck function and was determined in a two-step process.

Step 1: Conversion of level-1 DN values of a satellite imagery thermal infrared data to at-satellite

(ToA) Spectral Radiance Values

$$L_{\lambda,ToA} = M_L \times Q_{cal} + A_{L,} \qquad \text{…….3-22}$$

where ML is the multiplicative scaling factor for the given band. For Landsat, it is available in the image metadata file as RADIANCE_MULT_BAND_n, with n being the band;

AL is the radiance additive scaling factor for the given band. For Landsat it is available in the same file as RADIANCE_ADD_BAND_n, with n being the band and

Qcal is the level-1 pixel value stored as DN values in the image for Landsat it is available for both OLI and TIRS bands e.g., in Landsat 8.

Step 2: Involved the use of LλToA Image data from the equation above to calculate ToA in the K (Kelvin) unit by inverting Planck's radiation equation.

$$T_{ToA} = \frac{K_2}{\ln\left(\frac{K_1}{L_{\lambda ToA}} + 1\right)}, \qquad \text{.....3-23}$$

where Toa K1 and K2 are the thermal conversion constants for the given band. They are usually available in the image metadata file. For Landsat, they are available as K1_CONSTANT_BAND_n and K2_CONSTANT_BAND_n with n being the band.

The ToA brightness temperature in degrees Celsius(°C) units can be estimated by subtracting 273.15

in Equation above.

$$T_{ToA} = \frac{K_2}{\ln\left(\frac{K_1}{L_{\lambda ToA}} + 1\right)} - 273.15 \qquad \text{…….3-24}$$

These parameters were then used in equation 3-15 and used to calculate the land surface temperature. The result for these LST products was obtained from a model created using ARCGIS Model Builder in which all the parameters highlighted through equations 3-16 to 3-21 were been taken into account.

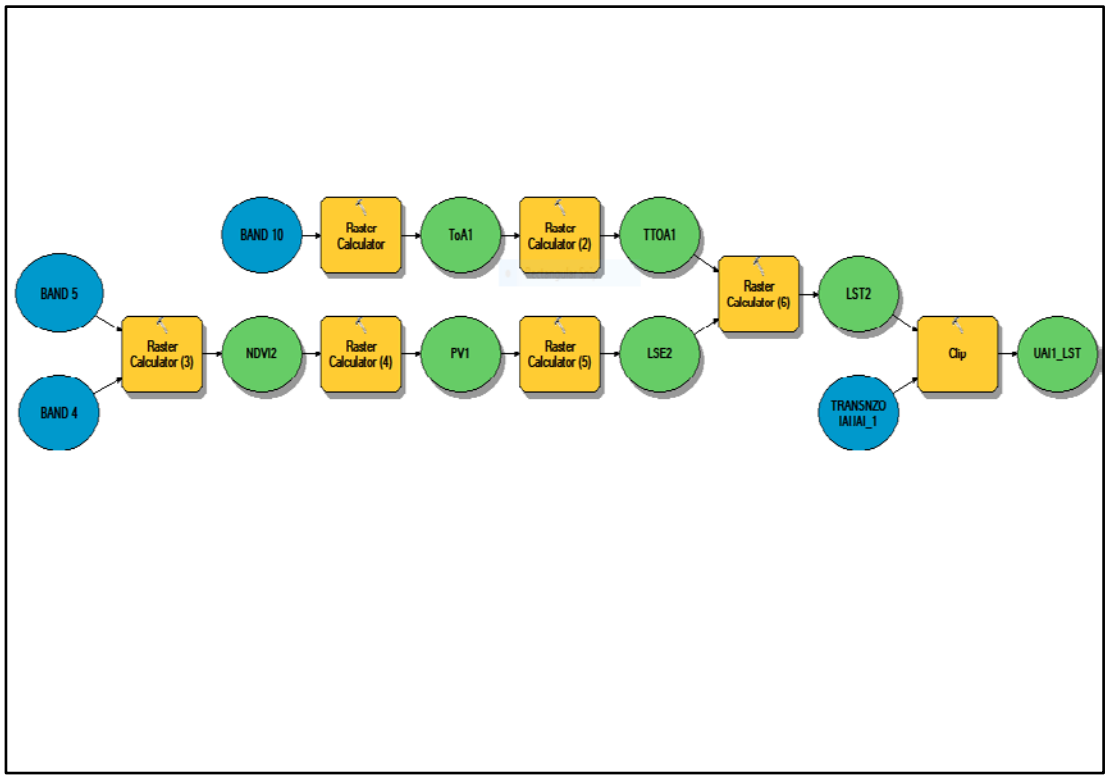


Figure 3.11: A model showing the process involved in the calculation of the land surface temperature.

3.4.1.7 Crop biomass

The study used nondestructive methods such as Spectral vegetation indices especially the Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI) and Soil Adjusted Vegetation Index (SAVI. The methodology illustrated in Figure 3.12 below was used to establish the crop biomass parameter

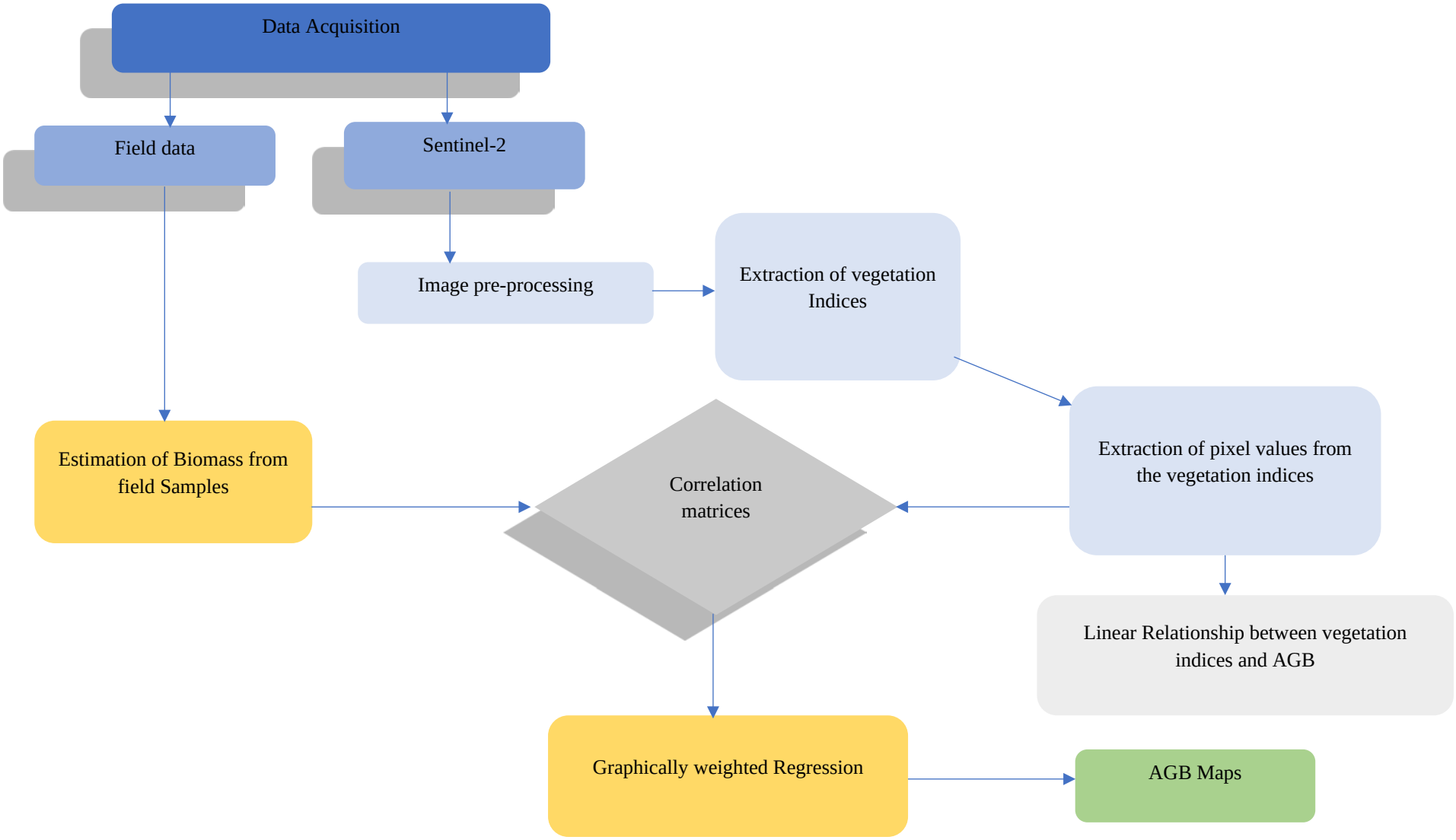


Figure 3.12 Methodology for biomass estimation

3.4.1.8 Soil Nutrients

1. Soil nitrogen

Methods based on remote sensing with spectral and thermal approaches were used as potential indicators to allow fast identification of soil nitrogen status. Satellite data observations in the visible and near-infrared spectral spectra provided information on leaf chlorophyll content, which permits the early detection of plant nutrient deficiency. The methodology used in the calculation of soil nitrogen is shown in Figure 3.13.

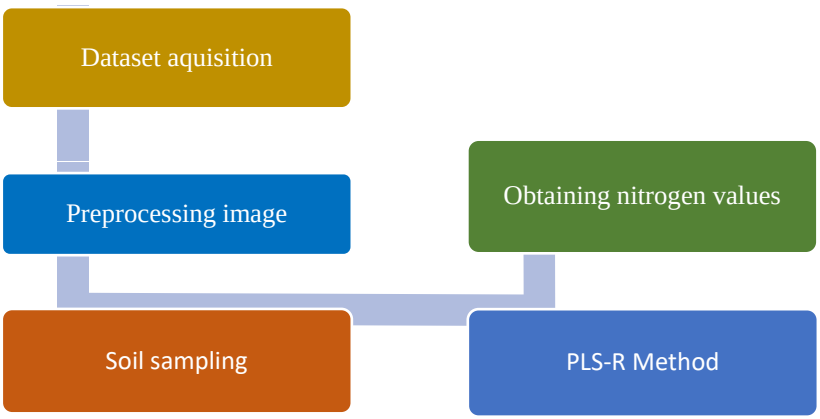


Figure 3.13: Methodology for extraction of soil Nitrogen

The satellite imagery data from Landsat 8 OLI/TIRS was downloaded, and the three bands needed for true-color imagery are: B4 is red (0.64–0.67 µm), B3 is green (0.53–0.59 µm), and B2 is blue (0.45–0.51 µm). After acquiring the image, atmospheric correction of the image was done, and cropping of the study area extent was established. Furthermore, the 250M resolution of soil nitrogen was obtained from the ISSRIC website, and using the ArcMap extraction tool from ArcMap, the soil values of random points on the map were obtained, including those obtained in the field, to generate enough data for the estimation of nutrients. MCRI2 and SAVI were used together with the blue band, red band, NIR band, SWIRR1 band, and SWIR2 band.

2. Extraction of organic carbon, phosphorus, and potassium from soil

Soil samples were generated from existing spatial soil nutrients datasets from the ISRIC website. These were used alongside field data collected during soil sampling. Table 3.5 provides information on the data used for sampling.

Table 3.5: Sources of soil nutrient data sets for use in sampling

Data	Measurement unit	Depth(cm)	Link
Total Phosphorus	Mg/kg	0 - 30	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/f72f5698-4a3d-4af2-ae12-a4b1cf151cec
Exchangeable Potassium	Cmolc/kg	0 - 20	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/48d2e636-6bda-4fa3-93c0-fb689fcf3340
Soil Organic Carbon	g/kg	0 - 5	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/9a66a37e-8a4e-463b-b83a-fd49049c323a

The raster files obtained were for potassium, phosphorus, and soil organic carbon at 250 m resolution. Field samples were used in conjunction with other generated random samples within the study area. Landsat 8 satellite imagery was used in this study. The Landsat images were subjected to radiometric and atmospheric corrections. The atmospheric correction was performed using the dark object subtraction method. From the satellite image, the normalized difference vegetation index (NDVI) and the land surface temperature were derived. The slope and elevation were also generated using a digital elevation model. pH, aluminum, iron, and cation exchange data were also downloaded from the ISRIC portal.

Table 3.6: Derived variable of CEC, extractable Al, and extractable iron

Data	Measurement unit	Depth(cm)	Link
CATION EXCHANGE	cmolc/kg	0 - 5	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/e0d921ff-5f7b-48e5-ae27-7c1515055e3b
EXTRACTABLE ALUMINIUM	mg/kg (ppm)	30	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/a36f7919-0d6e-4044-902c-64a74feade6b
EXTRACTABLE IRON	mg/kg (ppm)	30	https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/5cd5336c-2f45-4430-a9a8-312aa2095cb6

The organic carbon was investigated using elevation, slope, land surface temperature, NDVI, and pH, phosphorus was investigated using elevation, slope, land surface temperature, aluminum, iron, cation exchange, and pH. Potassium was investigated using cation exchange, land surface temperature, and pH. The equations based on the parameters selected were then used to predict soil nutrients within the study area. The ArcGIS model builder was used to create a model that incorporates these equations. Figures 3.14 and 3.15 below that shows flow chart of the methodology and the model developed in an ARCGIS environment for use in the estimation of NPK and organic carbon of the study area.

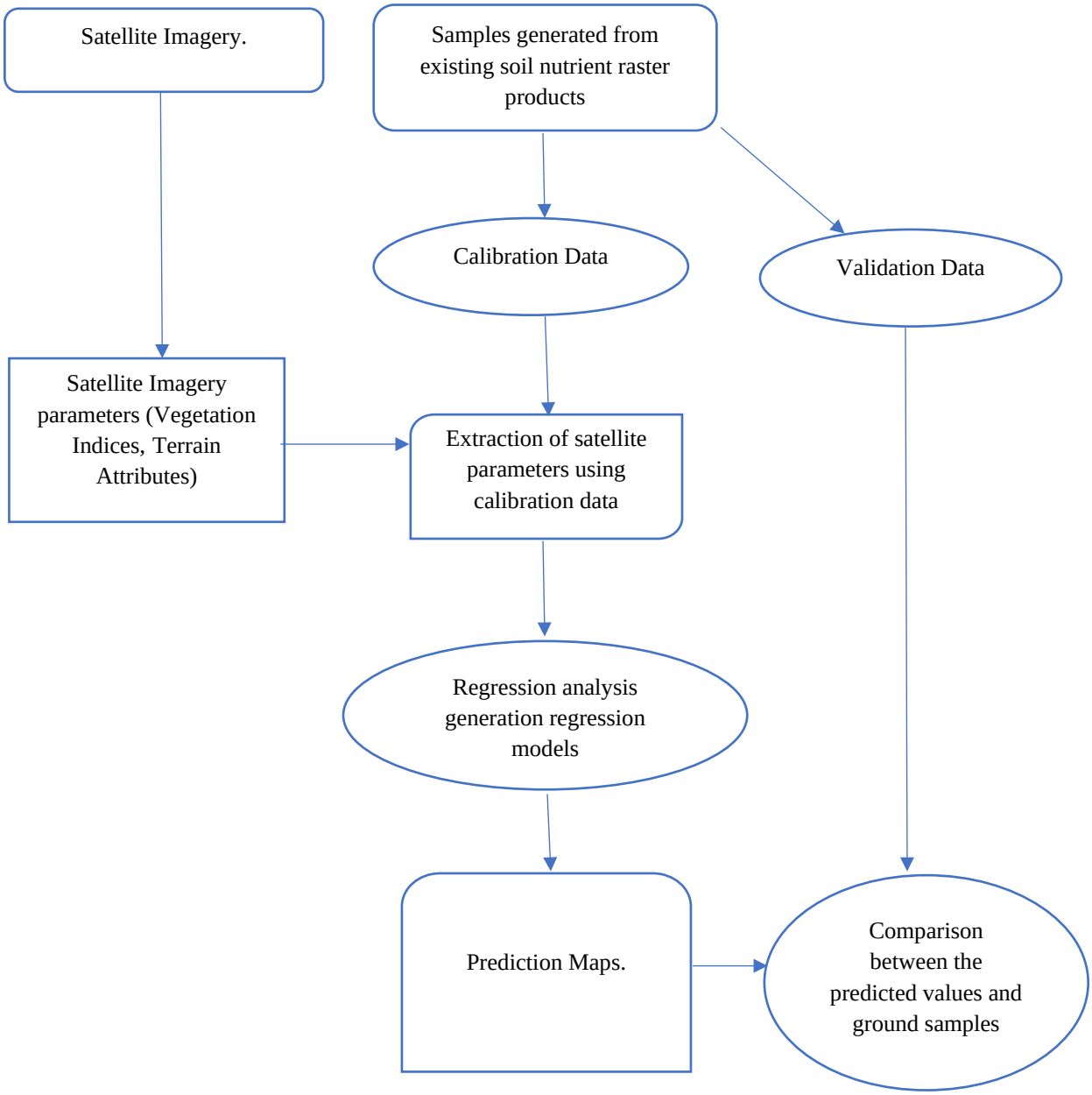


Figure 3.14: Methodology for extraction of organic carbon, phosphorus, and potassium from soil

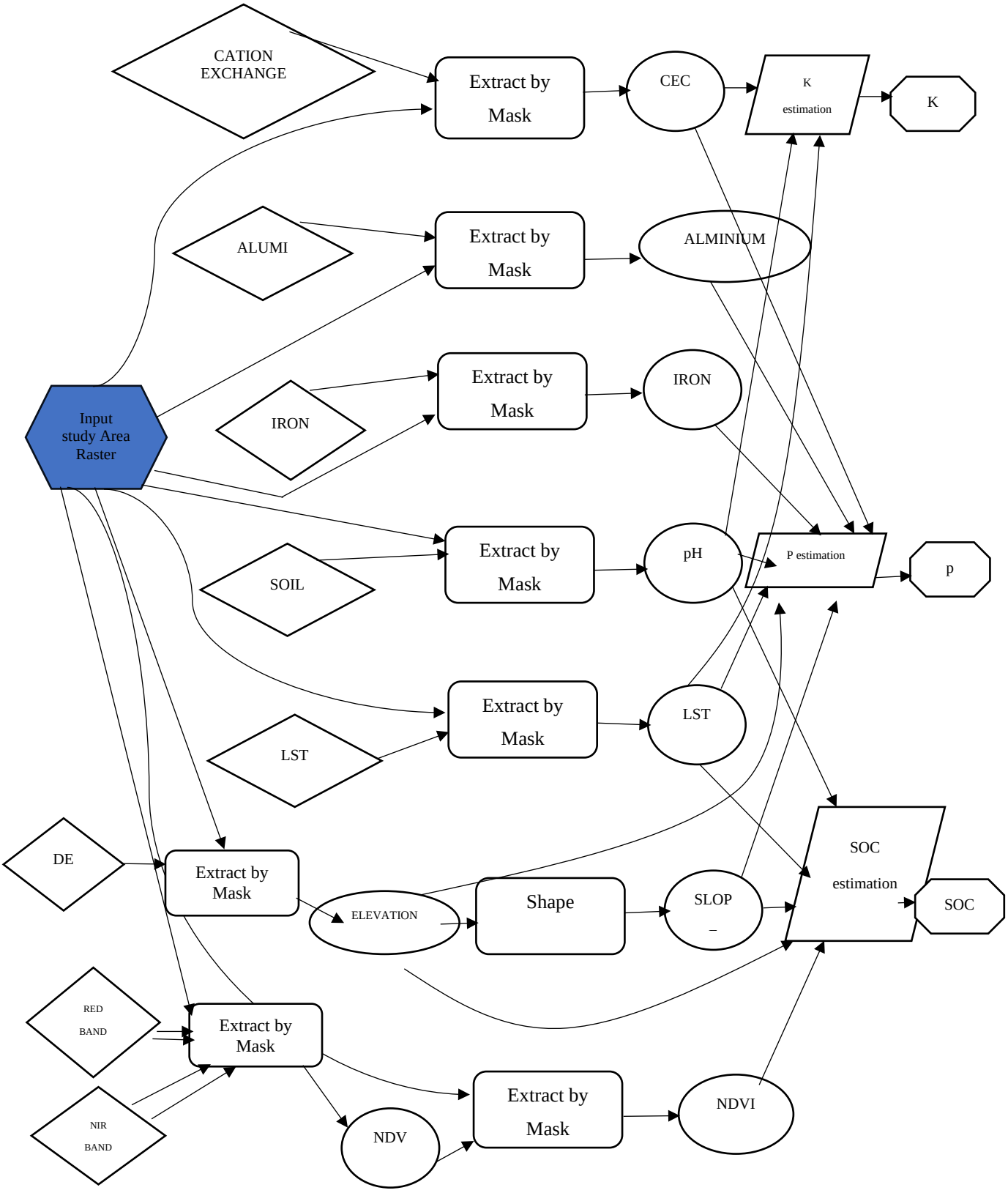


Figure 3.15: Model developed in ARCGIS environment for soil organic carbon, phosphorus, and potassium extraction

3.5 Calibration and Validation of the random forest machine learning yield estimation model

Elkan proposed a general approach to update the probability estimates for a binary outcome y to a population with a different unconditional event probability, termed the base rate. b = P(y = 1) is the base rate in the population in which the model has been developed and assume the availability of a model for which the probabilities P(y = 1|x) can be estimated for observations with characteristics x. To obtain updated probability estimates P′(y = 1|x) for observations of another population with base rate b′ = P′(y = 1), it is assumed that the change in base rate is the only difference between the two populations. In particular, the assumption is that the distribution of individual characteristics remains the same in both classes, that is, P(x|y = 0) = P′(x|y = 0) and P(x|y = 1) = P′(x|y = 1). Under these assumptions, P′(y = 1|x) can be expressed as a function of P(y = 1|x), b and b′

$$P'\left(y=1\middle|x\right)=\frac{b'P(y=1|x)-bb'P(y=1|x)}{b'P(y=1|x)+b-\mathrm{bP}(y=1|x)-bb'}. \qquad \text{……3-22}$$

3.5.1 Accuracy metrics that were utilized in the validation of the yield estimation model are

The root mean square error (RMSE), correlation factor (R), and relative mean absolute error (RMAE) are common accuracy metrics that were used in validating the accuracy of the GIS-based yield estimation model. The root mean square error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points they are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data are around the line of best fit. The correlation factor (R) measures the linear relationship between the predictions of the regression model and the real values. The mean absolute error (MAE) is the average of the estimation differences (in physical units). This metric is expressed as a percentage relative to the mean yield, being called

RMAE instead of MAE. The equation below shows how these metrics are calculated, where y is the real yield value, Ŷ represents the yield estimation, i is the number of samples, ȳ is the average of the real yield values, and Ŷ is the average of predictions:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad \text{……..3-23}$$

$$\text{RRSE}(\%) = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \times 100,$$

$$R = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}})^2}},$$

$$\text{RMAE}(\%) = \left(\frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{(n)(\overline{y})}\right) \times 100. \quad \textbf{……3-24}$$

3.6 GIS-based farm data management tool for sorghum and green grams

This study used the ARCGIS model builder to develop a crop data estimation tool that has a series of models that drive towards the final output, i.e., estimated crop area and crop signature range development, microclimatic mapping of the agroecological zone, and finally yield estimation using remote sensing data based on the data from the previous two steps, which form part of the entire model. The analytical part consists of three modules: crop area estimation module, crop signature module, agro-ecological sub-zonation module, and finally yield estimation module. The yield estimation module has three sub-modules: the meteorological module, the soil module, and the crop growth module.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Spatial distribution and spectral signatures of sorghum and green grams

4.1.1 Spatial estimation of cropping patterns in the study area under farm field conditions

The cropping pattern identified in the field for the Ikombe-Katangi study area was greengram, maize, mixed crop maize and beans (MB), mixed crop maize and cowpea (MC), mixed crop maize and greengram (MG), mixed crop maize and pigeon pea (MP), and vegetables (Figures 4.1 a and b, respectively) and Table 4.1.

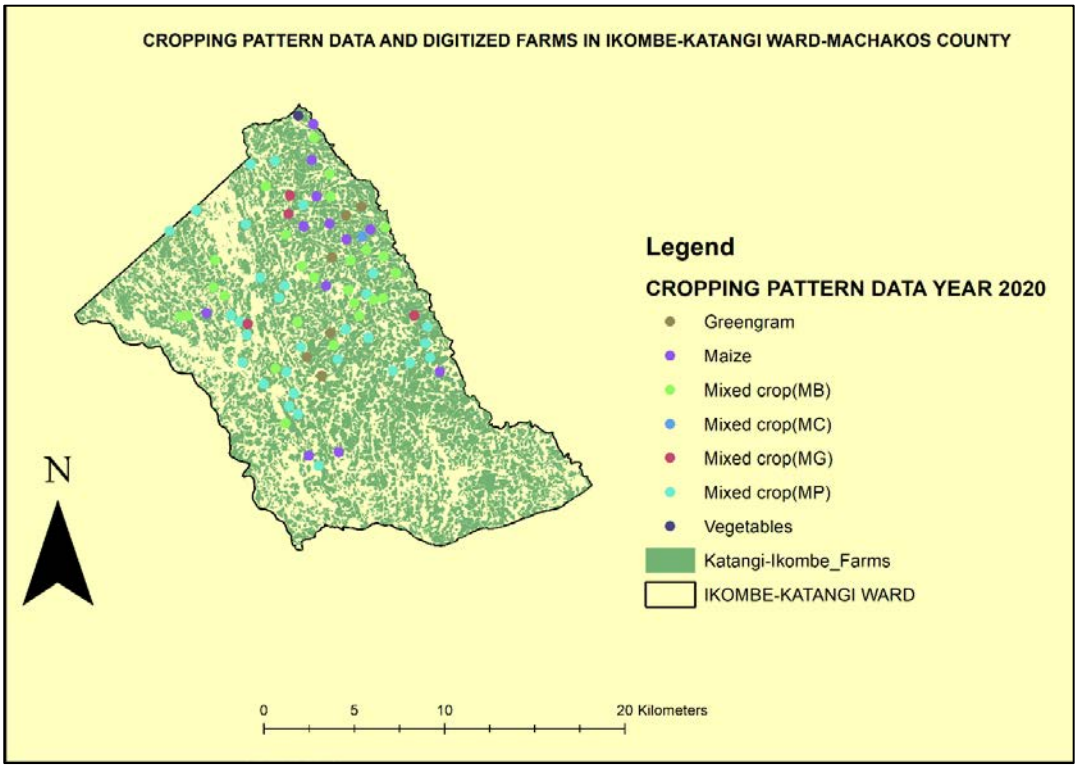


Figure 4.1a: Digitized farms with cropping pattern data for crop identification Ikombe-Katangi study area

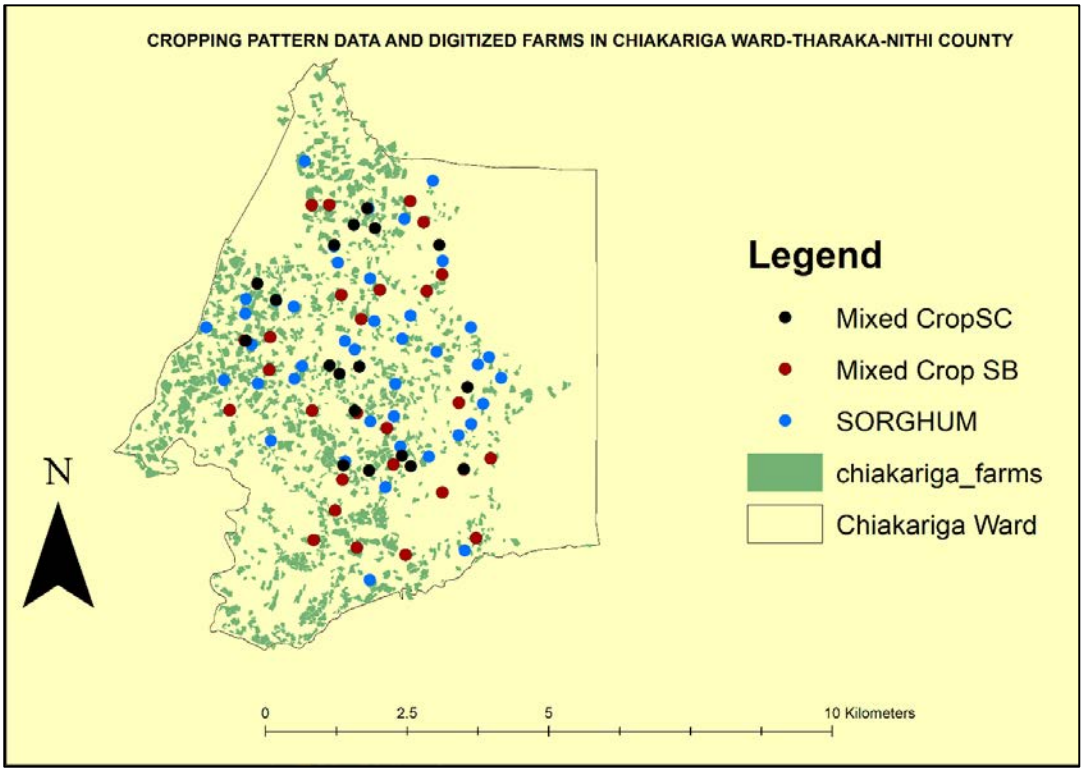


Figure 4.1b: Digitized farm with cropping pattern data for crop identification Chiakariga study area

Table 4.1: Cropping pattern data for Chiakariga and Ikombe- Katangi study area for year 2020,2018,2017,2015 and 2013.

FarmID	Farm Size	2020	2018	2017	2015	2013
1	2	Mixed crop (MP)	Mixed crop(MPB)	Mixed crop(MP)	Mixed crop(MC)	Mixed crop(MP)
2	3	Mixed crop (MP)	Maize	Maize	Mixed crop(MP)	Maize
3	4	Mixed crop (MP)	Mixed crop (MB)	Maize	Mixed crop(MP)	Mixed crop(MP)
4	3.5	Mixed crop (MP)	Maize	Maize	Mixed crop(MB)	Greengram
5	2	Mixed crop(MP)	Mixed crop(MP)	Mixed crop(MB)	Maize	Mixed crop(MP)
6	3	Mixed crop(MP)	Maize	Maize	Maize	Mixed crop(MB)
7	8	Mixed crop(MP)	Maize	Mixed crop(MB)	Maize	Mixed crop(MP)
8	3	Mixed crop(MP)	Mixed crop(MP)	Maize	Mixed crop(MB)	Mixed crop(MP)
9	5	Mixed crop(MP)	Maize	Mixed crop(MP)	Mixed crop(MB)	Mixed crop(MC)
10	5	Mixed crop(MP)	Maize	Maize	Mixed crop(MP)	Mixed crop(MB)
11	2	Mixed crop(MG)	Mixed crop(MB)	Mixed crop(MP)	Mixed crop(MP)	Mixed crop(MP)
12	2	Mixed crop(MP)	Mixed crop(MP)	Mixed crop(MPB)	Mixed crop(MP)	Mixed crop(MP)
13	2.5	Mixed crop(SC)	Mixed crop(SC)	Sorghum	Mixed crop(SB)	Mixed crop(SC)
14	2	Sorghum	Mixed crop (SC)	Sorghum	Mixed crop(SC)	Mixed crop(SB)
15	2	Sorghum	Mixed crop (SB)	Mixed crop(SC)	Mixed crop(SC)	Sorghum

The identified cropping pattern data has been presented in Table 4.1. This farmer survey was able to identify common cropping patterns for study areas within a 5-year period. Cropping intensity through multiple cropping is important for food security and the development of policy (Wu *et al*., 2018). As such, the need for correct and efficient cropping pattern maps is becoming increasingly important (Zhong *et al*. 2014). Furthermore, the spatial distribution and temporal dynamics of crop planting patterns are important in the understanding of climate change and the overall response to environmental issues emanating from agriculture (Wu *et al*., 2018). Establishing these patterns is key to mapping spectral signature libraries. Establishing the difference between the planted area and harvesting is critical to understanding the available food for the people. Establishing cropped areas in developing countries is important for addressing food security issues. In addition, Wu *et al*. (2018) indicated

that the spatial distribution and temporal dynamics of crop planting patterns have great implications for the overall understanding of agricultural production. Smallholder farms are usually subdivided into parcels whose boundaries are hard to distinguish and/or clearly delineate (Fritz *et al*., 2019). Spatial information on agricultural fields is important for the appropriate understanding, planning, and management of crop growth conditions. Graesser and Ramankutty (2017) detected cropland field parcels using Landsat imagery. The study showed success in the use of 30 m of Landsat satellite imagery in farm delineation and crop pattern identification and the results agrees with those of Yan and Roy (2014) who provided the capabilities of crop field extraction from multi-temporal web-enabled Landsat data. The study indicated that, through an understanding of crop phenologies, complex crop probability maps can be generated, which is in concurrence with the current study. Medium-spatial-resolution remote sensing data provides the spectral-temporal profiles of the crops of interest using a set of multispectral images acquired within a temporal period of time (Qiu *et al*., 2022). The current research findings are in agreement with those of Feyisa *et al*. (2020), who also used a combination of farmer surveys and remote sensing was found to improve cropping pattern data accuracy. A study by HAO *et al*. (2020) similarly pointed out the challenges of generating early-season crop type distribution at 30 m. The research by HAO *et al*. (2020), the maximum vegetative growth period was found to be good for cropping pattern identification. The use of single-date satellite data, as was the case for current research, is confirmed by Potgieter *et al*. (2021), who showed various situations when single-date satellite data could be used, i.e., areas with limited images due to cloud cover. Cloud cover affected the availability of multitemporal images for this research, so a single-date image was considered every year. This is also argued by Matvienko *et al*. (2020) in their research on single crop classification approaches. The research examined the possibility of using single-date satellite data. Limitations of Landsat data were pointed out in a study where Saini and Ghosh (2018) identified complex reflectance signals within and between fields to affect crop type determination.

To counter the challenge, the research used boundary delineations, as argued by Castillejo-González *et al*. (2009), where segmentation of images and object-based classification reduce noise that brings spectral noise. The thematic cropland results capture the key cropping patterns of the study area, and the detection of individual field parcels illustrates the broad range of field sizes across this vast area. The crop classification gives the accuracy of these methods in data estimation.

4.1.2 TOA Reflectance's calculations

The calculation of TOA reflectance for bands 2, 3, 4, 6, and 7, which are very important in the identification of cropped areas, was done, and the data is represented in Tables 4.2a and b for the Ikombe-Katangi and Chiakariga study areas. The top of the atmosphere reflectance values were different for each cropping pattern in the Ikombe-Katangi study area, as shown in Table 4.2a. Similarities in reflectance values were observed over the years for each of the cropping patterns. Table 4.2a Calculation of the top of the atmosphere reflectance for all the cropping patterns in bands 2, 3, 4, 6, and 7 for the five-year period in the Ikombe-Katangi study area.

Table 4.2a Calculation of the top of atmosphere reflectance for all the cropping patterns for the band 2,3,4,5,6 and 7 for the five-year period for Ikombe-Katangi

Year	Band Wavelength	Mixed Crop (MPB)	Mixed Crop (MG)	Mixed Crop (MB)	Maize	Green Gram	Mixed Crop (MP)	Mixed Crop (MC)	Mixed Crop (MPC)	Bean	Mixed Crop (MCB)
2013	0.455	0.08927	0.09586	0.09782	0.09714	0.09611	0.09523	0.09042	0.09455	0.10394	0.00000
2013	0.56	0.08397	0.09349	0.09682	0.09286	0.09467	0.09320	0.08738	0.09497	0.10117	0.00000
2013	0.655	0.11389	0.11269	0.12533	0.11367	0.12188	0.11541	0.11783	0.11799	0.12419	0.00000
2013	0.865	0.17718	0.19634	0.19709	0.17834	0.20016	0.18541	0.19077	0.18765	0.19208	0.00000
2013	1.62	0.24180	0.24705	0.27321	0.24897	0.26900	0.26013	0.26677	0.26864	0.28703	0.00000
2013	2.2	0.19331	0.19179	0.21507	0.19643	0.20815	0.20364	0.20649	0.20739	0.23309	0.00000
2015	0.455	0.08178	0.07869	0.07738	0.07785	0.07845	0.07767	0.07505	0.07827	0.00000	0.00000
2015	0.56	0.08142	0.07211	0.07469	0.07497	0.07489	0.07526	0.07280	0.07605	0.00000	0.00000
2015	0.655	0.08380	0.07236	0.08117	0.08121	0.07564	0.07716	0.06525	0.08031	0.00000	0.00000
2015	0.865	0.24936	0.19873	0.23230	0.22657	0.22178	0.23125	0.26128	0.22892	0.00000	0.00000
2015	1.62	0.22320	0.18761	0.20874	0.20457	0.19699	0.20628	0.17366	0.19619	0.00000	0.00000
2015	2.2	0.15126	0.12995	0.14266	0.14050	0.13043	0.13868	0.10051	0.13174	0.00000	0.00000
2017	0.455	0.07797	0.07567	0.07891	0.07801	0.07696	0.07651	0.07704	0.00000	0.00000	0.07734
2017	0.56	0.07738	0.07425	0.07842	0.07528	0.07466	0.07421	0.07032	0.00000	0.00000	0.07668
2017	0.655	0.09232	0.08937	0.08987	0.08499	0.08390	0.08719	0.07688	0.00000	0.00000	0.09117
2017	0.865	0.19938	0.19458	0.20790	0.19507	0.19822	0.19126	0.15944	0.00000	0.00000	0.20300
2017	1.62	0.23083	0.21366	0.22584	0.21146	0.19725	0.21912	0.19497	0.00000	0.00000	0.22754
2017	2.2	0.16924	0.15447	0.15943	0.15161	0.14430	0.15880	0.14508	0.00000	0.00000	0.16386
2018	0.455	0.08929	0.09017	0.09379	0.09033	0.09632	0.09112	0.09249	0.08673	0.08868	0.09468
2018	0.56	0.08767	0.09264	0.09586	0.08924	0.09672	0.09212	0.09578	0.08645	0.08932	0.09685
2018	0.655	0.10684	0.12070	0.12362	0.11328	0.11869	0.12156	0.12807	0.11046	0.11701	0.11856
2018	0.865	0.20444	0.21914	0.22083	0.20093	0.20667	0.21193	0.22814	0.20931	0.21768	0.21566
2018	1.62	0.23805	0.26873	0.29415	0.25914	0.26891	0.27938	0.29846	0.24534	0.26241	0.29436
2018	2.2	0.17199	0.19829	0.22210	0.19884	0.20345	0.21170	0.23097	0.18363	0.19895	0.21677
2020	0.455	0.00000	0.07989	0.07765	0.07880	0.08206	0.07922	0.00000	0.00000	0.00000	0.00000
2020	0.56	0.00000	0.07697	0.07615	0.07766	0.07916	0.07814	0.00000	0.00000	0.00000	0.00000
2020	0.655	0.00000	0.07061	0.06397	0.06949	0.07047	0.06976	0.00000	0.00000	0.00000	0.00000
2020	0.865	0.00000	0.24227	0.27798	0.26834	0.24387	0.26933	0.00000	0.00000	0.00000	0.00000
2020	1.62	0.00000	0.18978	0.18944	0.19581	0.19269	0.19582	0.00000	0.00000	0.00000	0.00000
2020	2.2	0.00000	0.11448	0.10414	0.11376	0.11354	0.11367	0.00000	0.00000	0.00000	0.00000

Spectral signatures, in general, can be defined as characteristics of surface objects in the transmission, absorption, and reflection of electromagnetic radiation. The spectral signature is expected to be stable and unique for a given surface material. Spectral signatures have been used for a long time for object detection and classification as

similarly reported by Duong *et al*. (2014). The ability to distinguish between the green-up and senescence stages is especially important for the interpretation of satellite data over regions where multiple cropping is common as emphasized by Nguy-Robertson *et al*. (2013). Spectral spaces, that is, relationships between reflectances in different spectral bands, can provide additional information, from that gained from vegetation indices, about the phenological and physiological status of vegetation (Nguy-Robertson *et al*., 2013). The reflectance values that determine the spectral signature libraries for sorghum were observed for all bands across the different cropping patterns and years, as shown in Table 4.2b. The TOA spectral reflectances were different for the different cropping patterns but showed similarities across the years for each cropping. Several studies have supported the use of spectral reflectance in retrieving the properties of vegetation. For example, by using spectral signatures, useful information can also be retrieved by analyzing the directional reflectance properties of vegetation (Wardlow *et al*., 2007). Reflectance, shown by the spectral signature, is connected to the absorption and transmission of each wavelength. This is an indication of plant performance in real life or in an experiment (Garriga *et al*., 2014). Shelestov *et al*. (2017) indicate that images acquired at different dates during the crop growth period are usually required to discriminate against certain crop types. Further, Shelestov *et al*. (2017) indicate that issues such as (i) nonuniformity of coverage of ground truth data and satellite scenes; (ii) seasonal differentiation of crop groups; are critical when handling multitemporal satellite images for cropped areas. The temporal resolution of satellite data is also critical for crop discrimination and mapping due to the dynamic character of agricultural systems (Lobell and Gourdji, 2012).

Table 4.2b Calculation of the top of atmosphere reflectance for all the cropping patterns for the band 2,3,4,5,6 and 7 for the five-year period for Chiakariga

	Crop	BAND2	BAND3	BAND 4	BAND 5	BAND 6	BAND 7
Year	Wavelength	0.455	0.56	0.655	0.865	1.62	2.2
2020	SORGHUM	0.32112577778	0.50665555556	0.58084655556	0.25338822222	0.34301233333	0.39544044444
2018	SORGHUM	0.28889400000	0.45267160000	0.47863480000	0.16826400000	0.16191500000	0.15071280000
2017	SORGHUM	0.29937616667	0.47473366667	0.50826100000	0.17148000000	0.16677800000	0.15194983333
2015	SORGHUM	0.25767883333	0.42981683333	0.51798316667	0.14225700000	0.15863100000	0.14692433333
2013	SORGHUM	0.23006400000	0.38210860000	0.47235073333	0.13865183333	0.14739180000	0.14354920000
2020	Mixed Crop SB	0.22082218333	0.37148566667	0.43560933333	0.11779908333	0.13447933333	0.13753716667
2018	Mixed Crop SB	0.12795785714	0.27312600000	0.52125757143	0.08368161429	0.12599942857	0.12923385714
2017	Mixed Crop SB	0.12287225000	0.26344275000	0.50516712500	0.08128298750	0.12502712500	0.12960975000
2015	Mixed Crop SB	0.21254851667	0.34515845000	0.47880116667	0.15284408333	0.17418625000	0.17337833333
2020	Mixed CropSC	0.29151820000	0.42769120000	0.41173240000	0.17652380000	0.15276240000	0.15094820000
2018	Mixed CropSC	0.30147633333	0.43665966667	0.37911333333	0.17603233333	0.16082333333	0.15890800000
2017	Mixed CropSC	0.26061300000	0.39292250000	0.39609050000	0.14095575000	0.14174325000	0.14496275000
2015	Mixed CropSC	0.21372000000	0.35873466667	0.48198600000	0.12918270000	0.13910666667	0.13843733333
2013	Mixed CropSC	0.19384900000	0.34371033333	0.48985600000	0.10693466667	0.13756033333	0.13874900000

This study used TOA reflectance, which enhances the identification of cropped areas, as well as development spectral reflectance libraries (Tables 4.2 a and b) and Figure 4.2. The study results are similar to those of from the research of Estévez *et al*. (2022), who underscored the need to rectify satellite data for atmospheric issues. The calculation of reflectance is crucial when creating multitemporal mosaics, as it largely removes variations between these images due to sensor differences, Earth-sun

distance, and solar zenith angle (caused by different scene dates, overpass time, and latitude differences) (Bruce and Hilbert, 2004). The conversion from DN to TOA reflectance has been applied in many studies. According to Mohan (1998), it is the most crucial step in producing 'accurate' spectral reflectance. Other studies, such as those by Fang and Liang (2003), have confirmed that the use of surface reflectance and TOA produced validly accurate results in the estimation of the LAI index. This has provided evidence that in some types of indices, such as LAI, atmospheric errors do not affect resultant outputs. Several studies have in the past advocated for the use of TOA to avoid potential errors that can result due to subsequent retrieval processes (Estévez *et al*., 2022). Analysis of atmospheric correction influence on spectral reflectance values has been researched by Liang *et al*. (2001) and Rumora *et al*. (2021), or by using just one classification method (Thenkabail and Wu, 2012). This particular step in the study ensured all the vegetation indices used in the crop yield estimation would be accurate. This part of the study looks at how spectral resolution, which is measured by green gram spectral signatures, impacts the amount and type of thematic information that can be gathered from a satellite scene, specifically how well it can tell the difference between plants. This argument is in accordance with those of Jensen (2000), who defined spectral resolution as the width of the spectral intervals (bands) used. A calculation of TOA reflectance increased the probability of identification of cropping patterns in the study area, especially with the focus of training data that had identified all possible cropping patterns. The current findings are in agreement with those of Hansen and Indeje (2004), who confirm that the identification of crops on agricultural surfaces was the best performed by multitemporal classification and the use of spectral reflectance, just as was the case for this particular study. Blickensdorfer *et al*. (2022) found that changes in meteorological variables make crop differences between regions bigger. To get around this problem, this study focused on agroecological zones where meteorological data is very similar. In further research, micro-agroecological factors were further examined for micro-climates to reduce these crop variations.

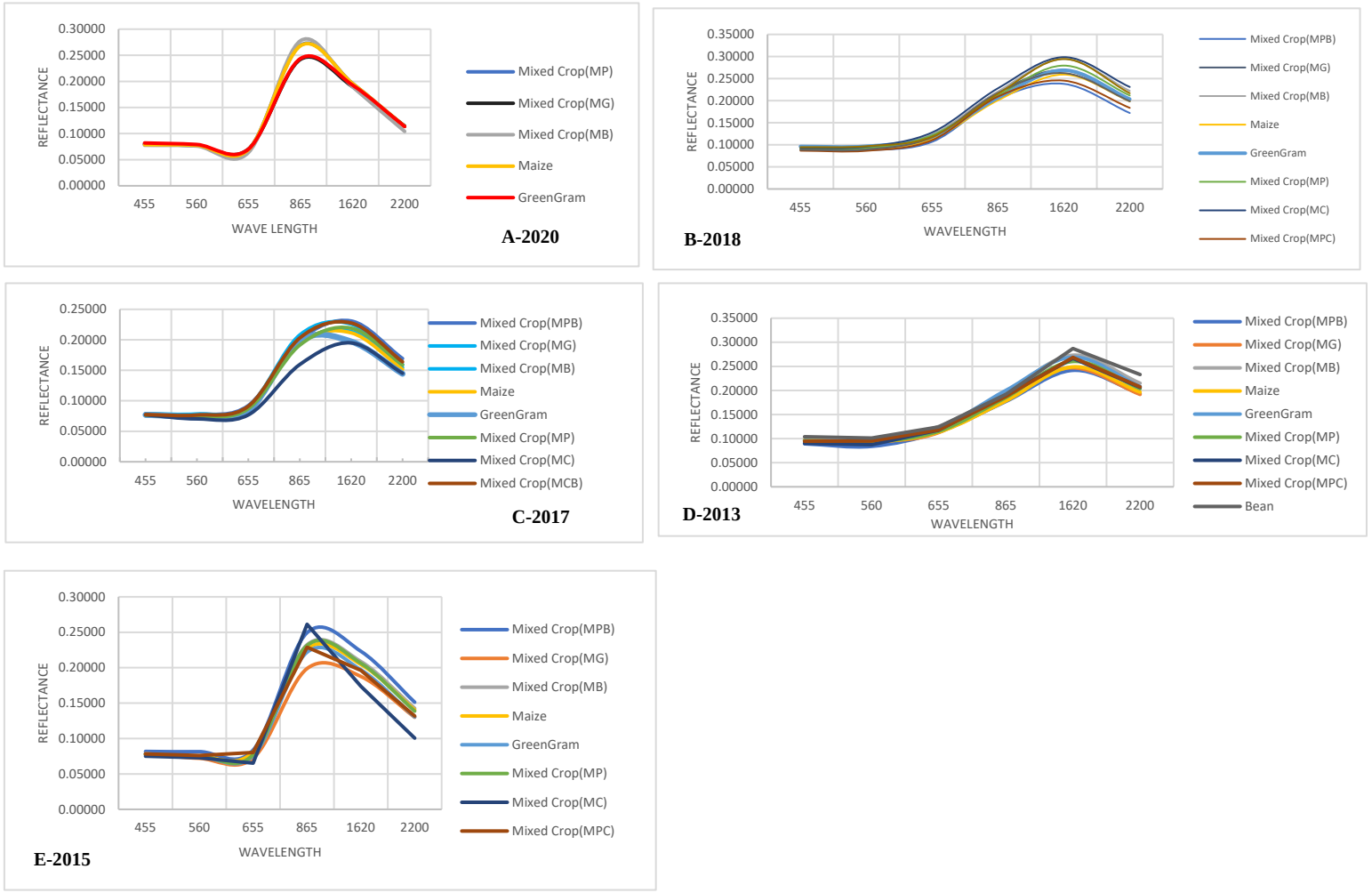


Figure 4.2: Spectral reflectance for Ikombe-Katangi study area for OND rains season

Separation of different classes and development of unique spectral libraries for each of the cropping patterns was achieved in this study, as shown in Figures 4.2 and 4.3 across different bands for sorghum and green gram. The spectral reflectance curves of different crops and crop mixtures were evident for all bands analyzed (2, 3, 4, 6, and 7). The spectral reflectance for the Ikombe-Katangi study area for the year 2018 is shown in Figure 4.2b. The ability to discriminate between crop types and crop mixtures is more evident at Bands 4, 5, 6, and 7. This is similar for spectral reflectance for the years 2020, 2017, 2015, and 2013, as shown in figures 4.2 a, c, d, and e. These findings agree with those of Ryerson *et al*. (1997) ; Ishimwe (2014) ; Nawaz *et al*. (2010), which suggest that the basis for separating one crop from another in remote sensing technology is the supposition that each crop species has a unique appearance or spectral signature. Similarly, the spectral signature of sorghum gave a positive result at bands 4, 6, and 7, as shown in figures 4.3 a, b, c, d, and e. The common practice for remote sensing is the utilization of red, NIR, and SWIR bands in vegetable identification (Wardlow and Egbert, 2010). Evidently, the current research was able to confirm this through the spectral reflectance discrimination being more evident in bands 4, 6, and 7, which are NIR, NIR, and SWIR. The ability of Landsat satellite data to enable crop type discrimination is reaffirmed by Shoko *et al*. (2018), who used Sentinel 2 and Landsat 8 data for spatial variation of C_3 and C_4 grasses. What is more important is the ability to differentiate grass species using Landsat data (Shoko *et al*., 2018). This confirms the research capabilities in crop type and crop pattern identification through spectral signature libraries for sorghum and green gram in the Kathaana and Ikombe-Katangi in the current study areas.

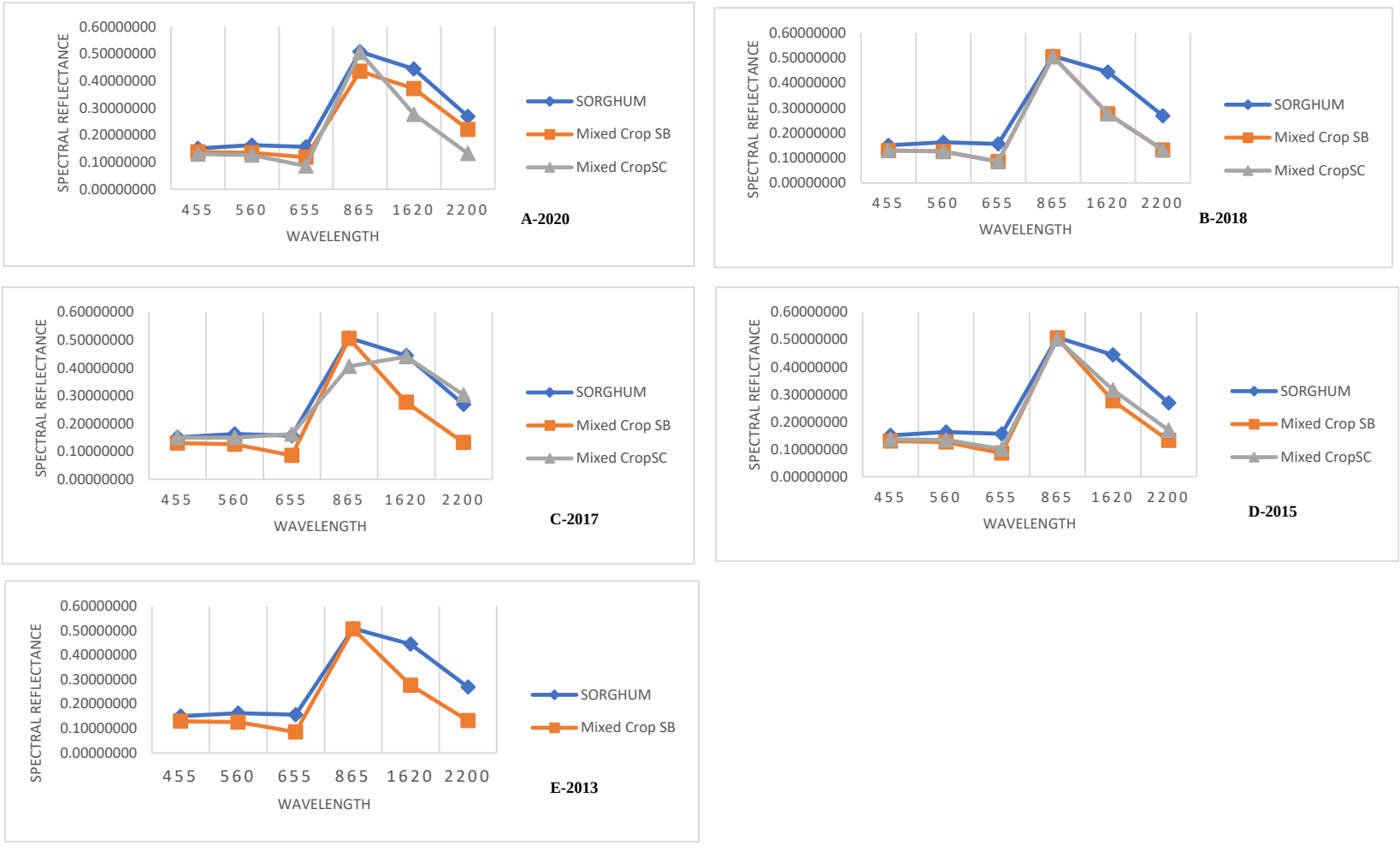


Figure 4.3: Spectral reflectance for Ikombe-Katangi study area OND rains season

The consideration of crop phenology while executing crop discrimination was also considered, where the research focused on the vegetative stage of growth, and this would give results superior to mono-temporal classifications as recommended by Kyei-Mensah *et al*. 2019; Clevers *et al*., 1994; Waldhoff *et al*., 2017). This is further confirmed by Shen *et al*. (2009), where variability in cropland is reduced through satellite data coverage of one key phenological phases of the crop, since surface reflectance changes with the growth stages of the crop. This particular approach improves the homogeneity of the environment by taking care of several factors that cause a heterogeneous nature in the identification of crops using satellite imagery. The use of cost-effective remote sensing data can lead to the development of a crop type database for the spatial-temporal identification of crop sequences (Waldhoff *et al*., 2017).

4.1.3 Assessment of the spectral reflectance discrimination for Ikombe-Katangi and Chiakariga study area

The proximity matrix using Euclidian distance was calculated and is represented in Tables 4.3 a and b for the Ikombe-Katangi and Chiakariga study areas. The table result showcases the levels of dissimilarity among the various cropping patterns for each of the TOA reflectance's.

Table 4.3a Euclidean distance proximity matrix for all cropping patterns identified in Ikombe-Katangi study area for all years and all reflectance bands

Proximity matrix (Euclidean distance):										
	Mixed Crop(MPB)	Mixed Crop(MG)	Mixed Crop(MB)	Maize	GreenGram	Mixed Crop(MP)	Mixed Crop(MC)	Mixed Crop(MPC)	Bean	Mixed Crop(MCB)
Mixed Crop(MPB)	**0**	**0.259**	**0.288**	**0.278**	**0.257**	**0.280**	**0.059**	**0.221**	**0.344**	**0.337**
Mixed Crop(MG)	**0.259**	**0**	**0.054**	**0.048**	**0.029**	**0.045**	**0.263**	**0.333**	**0.393**	**0.400**
Mixed Crop(MB)	**0.288**	**0.054**	**0**	**0.037**	**0.042**	**0.027**	**0.292**	**0.365**	**0.440**	**0.443**
Maize	**0.278**	**0.048**	**0.037**	**0**	**0.036**	**0.017**	**0.284**	**0.350**	**0.425**	**0.424**
GreenGram	**0.257**	**0.029**	**0.042**	**0.036**	**0**	**0.033**	**0.261**	**0.333**	**0.407**	**0.418**
Mixed Crop(MP)	**0.280**	**0.045**	**0.027**	**0.017**	**0.033**	**0**	**0.283**	**0.349**	**0.426**	**0.430**
Mixed Crop(MC)	**0.059**	**0.263**	**0.292**	**0.284**	**0.261**	**0.283**	**0**	**0.182**	**0.323**	**0.354**
Mixed Crop(MPC)	**0.221**	**0.333**	**0.365**	**0.350**	**0.333**	**0.349**	**0.182**	**0**	**0.243**	**0.397**
Bean	**0.344**	**0.393**	**0.440**	**0.425**	**0.407**	**0.426**	**0.323**	**0.243**	**0**	**0.319**
Mixed Crop(MCB)	**0.337**	**0.400**	**0.443**	**0.424**	**0.418**	**0.430**	**0.354**	**0.397**	**0.319**	**0**

The proximity matrix using Euclidian distance was calculated for three cropping patterns identified and surveyed in the Chiakariga study area. High levels of dissimilarities were identified across the cropping pattern reflectance values. The results are presented in Table 4.3b below.

Table 4.3b Euclidean distance proximity matrix for all cropping patterns identified in the Chiakariga study area for all years and reflectance bands

Proximity matrix (Euclidean distance):			
Cropping pattern	Mixed Crop(SC)	Mixed Crop(SB)	Sorghum
Mixed Crop(SC)	**0**	**0.364**	**0.386**
Mixed Crop(SB)	**0.364**	**0**	**0.079**
Sorghum	**0.386**	**0.079**	**0**

The analysis looked at the levels of similarities at a dissimilarity threshold of 0.95 as displayed in Table 4.4aandb. High levels of dissimilarities were identified across the cropping pattern reflectance values. In table 4.4a the dissimilarities of cropping

patterns of Ikombe-Katangi study area at dissimilarity threshold of 0.95 have been displayed.

Table 4.4a: List of similar objects at a dissimilarity threshold of 0.95 for Ikombe Katangi study area for all years and reflectance bands

List of similar objects(Dissimilarity threshold = 0.95):		
Object1	**Object2**	**Dissimilarity**
Mixed Crop (MPB)	Mixed Crop (MG)	0.259
Mixed Crop (MPB)	Mixed Crop (MB)	0.288
Mixed Crop (MPB)	Maize	0.278
Mixed Crop (MPB)	GreenGram	0.257
Mixed Crop (MPB)	Mixed Crop (MP)	0.280
Mixed Crop (MPB)	Mixed Crop (MC)	0.059
Mixed Crop (MPB)	Mixed Crop (MPC)	0.221
Mixed Crop (MPB)	Bean	0.344
Mixed Crop (MPB)	Mixed Crop (MCB)	0.337
Mixed Crop (MG)	Mixed Crop (MB)	0.054
Mixed Crop (MG)	Maize	0.048
Mixed Crop (MG)	GreenGram	0.029
Mixed Crop (MG)	Mixed Crop (MP)	0.045
Mixed Crop (MG)	Mixed Crop (MC)	0.263
Mixed Crop (MG)	Mixed Crop (MPC)	0.333
Mixed Crop (MG)	Bean	0.393
Mixed Crop (MG)	Mixed Crop (MCB)	0.400
Mixed Crop (MB)	Maize	0.037
Mixed Crop (MB)	GreenGram	0.042
Mixed Crop (MB)	Mixed Crop (MP)	0.027
Mixed Crop (MB)	Mixed Crop (MC)	0.292

List of similar objects(Dissimilarity threshold = 0.95):		
Mixed Crop(MB)	Mixed Crop(MPC)	0.365
Mixed Crop (MB)	Bean	0.440
Mixed Crop (MB)	Mixed Crop (MCB)	0.443
Maize	Green Gram	0.036
Maize	Mixed Crop (MP)	0.017
Maize	Mixed Crop (MC)	0.284
Maize	Mixed Crop (MPC)	0.350
Maize	Bean	0.425
Maize	Mixed Crop (MCB)	0.424
Green Gram	Mixed Crop (MP)	0.033
Green Gram	Mixed Crop (MC)	0.261
Green Gram	Mixed Crop (MPC)	0.333
Green Gram	Bean	0.407
Green Gram	Mixed Crop (MCB)	0.418
Mixed Crop (MP)	Mixed Crop (MC)	0.283
Mixed Crop (MP)	Mixed Crop (MPC)	0.349
Mixed Crop (MP)	Bean	0.426
Mixed Crop (MP)	Mixed Crop (MCB)	0.430
Mixed Crop (MC)	Mixed Crop (MPC)	0.182
Mixed Crop (MC)	Bean	0.323
Mixed Crop (MC)	Mixed Crop (MCB)	0.354
Mixed Crop (MPC)	Bean	0.243
Mixed Crop (MPC)	Mixed Crop (MCB)	0.397
Bean	Mixed Crop (MCB)	0.319

Levels of dissimilarities were analysed and identified across the cropping pattern reflectance values for Chiakariga study area. In table 4.4b the dissimilarities of cropping patterns of Chiakariga study area at dissimilarity threshold of 0.95 have been displayed.

Table 4.4b: List of similar objects at a dissimilarity threshold of 0.95 for Chiakariga study area for all years and reflectance bands

List of similar objects (Dissimilarity threshold = 0.95):		
Object1	**Object2**	**Dissimilarity**
Mixed Crop (SC)	Mixed Crop (SB)	0.047811
Mixed Crop (SC)	Mixed Crop (SB)	0.366819
Mixed Crop (SB)	Sorghum	0.49648
Mixed Crop (SB)	Mixed Crop (SC)	0.607394
Sorghum	Mixed Crop (SB)	0.664108
Sorghum	Mixed Crop (SC)	0.384362

Spectral signature matching algorithms have previously been used for vegetation species discrimination (Padma and Sanjeevi, 2014). The differentiation of the different crops was possible, as indicated by the spectral similarity assessment of the study. Another study by Ahram *et al*. (2015), also evaluated the use of various methods of spectral similarity assessment. The study highlighted the use of non-hybrid methods in the discrimination of objects with a high level of similarity. In this case, the Euclidean distance was selected for discrimination between the crop patterns. The method is deterministic and looks at the shape of the wave length (Vishnu *et al*., 2013). Vajsová *et al*. (2020) used the Pearson correlation coefficient to check similarities between the NDVI time series and Sentinel 2 data. Further, Padma and Sanjeevi (2014) indicate that Euclidian distance focuses on the amplitude of two spectra, resulting in more accurate results than other methods of spectral similarity assessment.

An in-depth analysis of the spectral signatures associated with green gram for all bands based on Figures 4.12 to 4.17 for green grams and Figures 4.18 to 4.23 gives an indication of the possibilities of having spectral signature ranges for the identification of crops or cropping patterns in the field. The spectral reflectance ranged from 0.07696-0.09632, 0.07466-0.09467, 0.0704047-0.12188, 0.19822-0.24387, 0.19269-0.26900, and 0.11354-0.20815 for bands two, three, four, five, six, and

seven, respectively, for green gram. In addition, the spectral reflectance for sorghum for bands two, three, four, five, six, and seven was 0.153-0.154, 0.163-0.167, 0.155-0.156, 0.533-0.54, 0.433-0.48, and 0.26-0.28, respectively. The graphs Figure 4.12 to 4.17 represent the spectral reflectance ranges for green gram for all bands evaluated in the study for a period of 5 years, while Figure 4.18 to 4.23 represent sorghum for all bands evaluated in the study for the same period of time. Crop spectral features, which are highly related to leaf pigment content, leaf water content, and crop canopy structure, are important information for crop classification (Hu *et al*., 2017; Jiang *et al*., 2006).

Further analysis of different bands across different years for green gram and sorghum was done and is represented from Figure 4.12 to 4.23. Significantly, the variations across the years are not far apart. Indicating that crop spectral signatures extracted for one season in a particular vegetative stage can relatively remain within a certain range. This is also confirmed by Jamal-Eddine *et al*. (2018) and Nguy-Robertson *et al*. (2013) studies, which indicate changes are more common across different crop growth cycles with minimal variability within certain crop growth stages.

This was to ensure that unique spectral signatures could be established for each cropping pattern. In the study by Ahram *et al*. (2015), who highlighted the use of nonhybrid methods such as Euclidean in the discrimination of objects with a high level of similarity. In this case, the Euclidean was selected for discrimination between the crop patterns in the study. The method is deterministic and looks at the shape of the wave length (Vishnu *et al*., 2013).

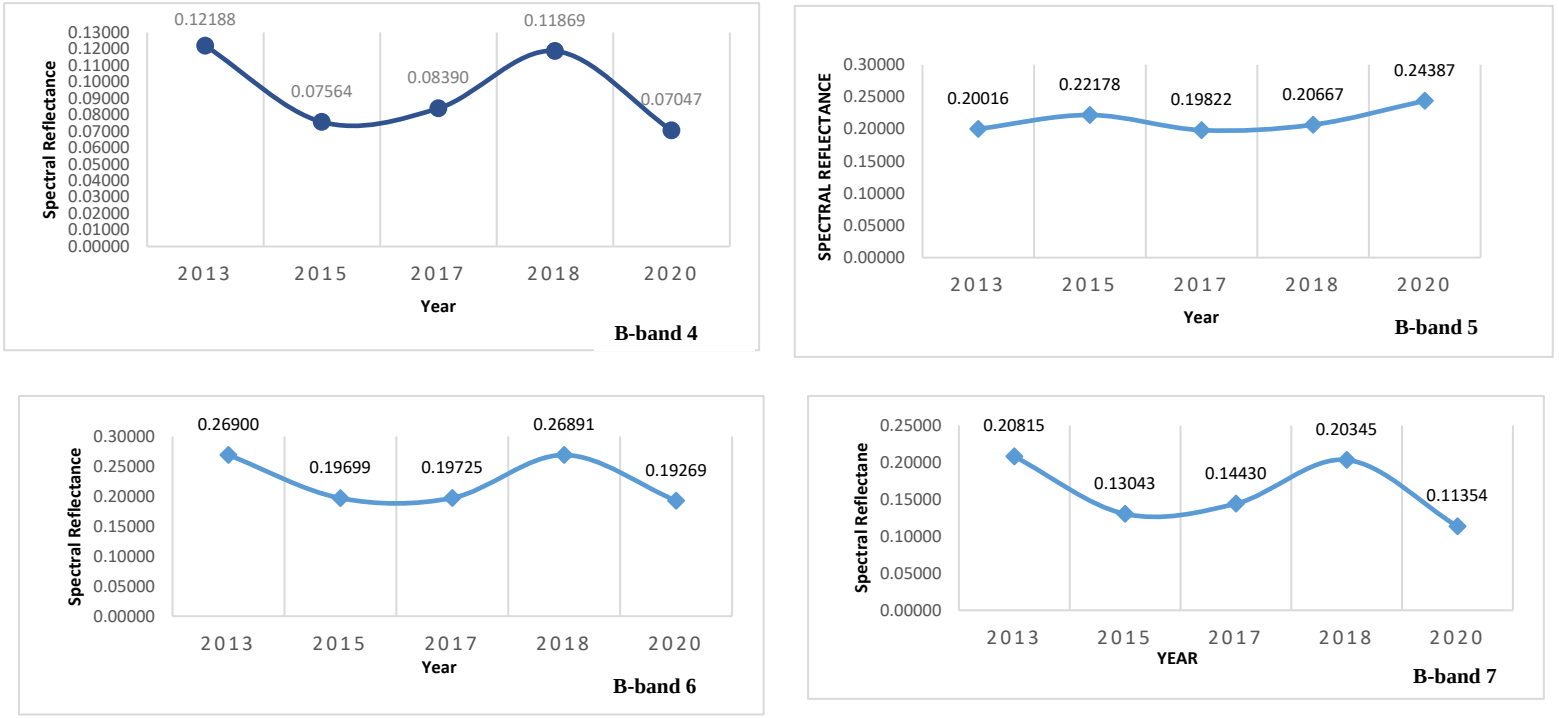


Figure 4.4: Green gram spectral reflectance signature range per band (4.5.6 and 7)

The spectral signature ranges for green gram were evaluated and extracted for bands 4, 5, 6, and 7 for the five years for the vegetative stage of growth in agro-ecological zones IV and V for the OND rain season, as shown in Figure 4.4. The spectral signature ranges for sorghum were evaluated and extracted for Bands 4, 5, 6, and 7 for the five years for the vegetative stage of growth in agro-ecological zones IV and V for the OND rain season, as shown in Figure 4.5.

October, November, and December are rainy season only. The similarities in the sigmoid curve of the TOA reflectance across the years are an indication that spectral signature libraries can be established through defined ranges for particular cropping stages, seasons, and agroecological zones. Such a similar finding with maize was previously reported by Kumar *et al*. (2021).

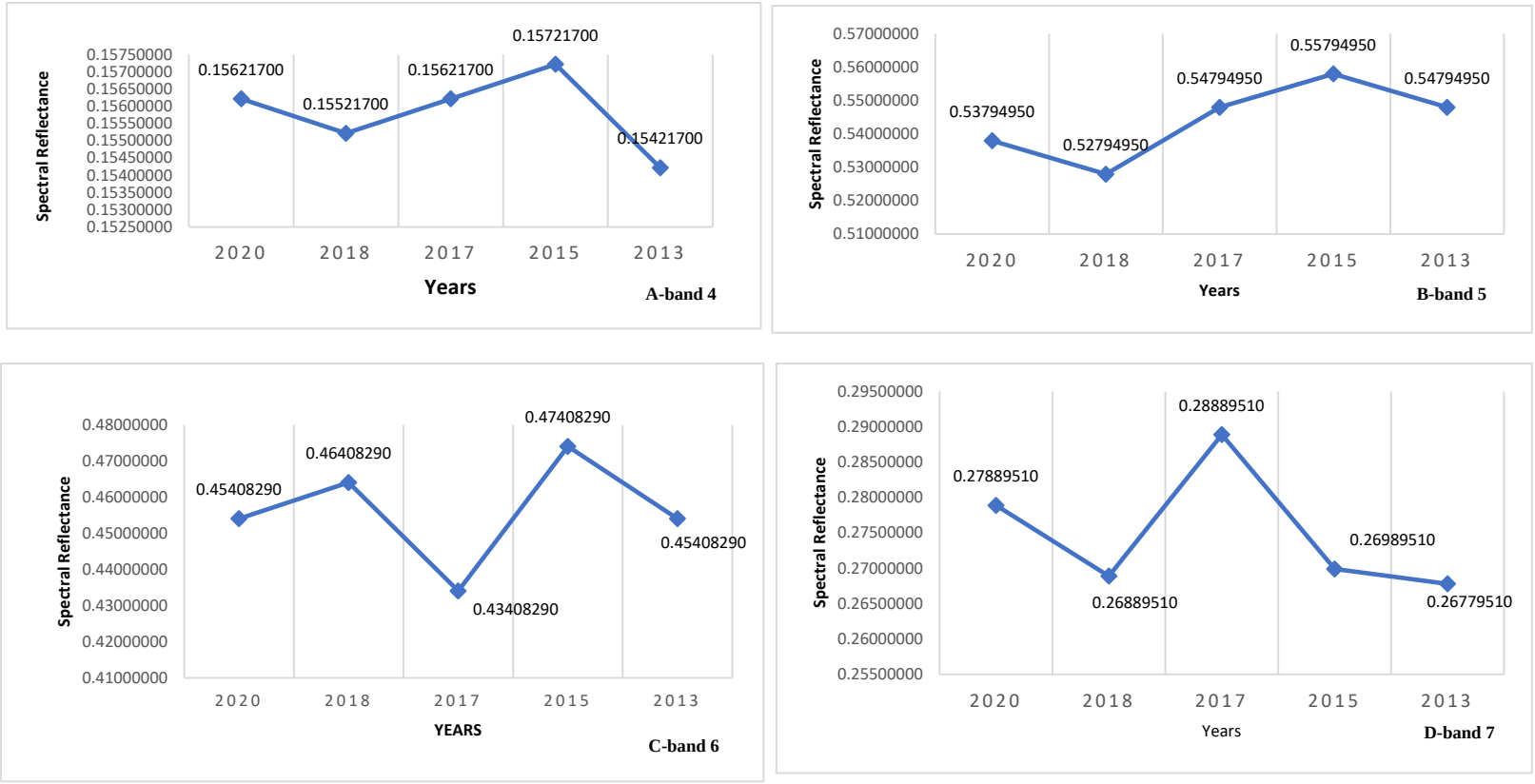


Figure 4.5: Sorghum spectral signature range for band (4,5,6 and 7)

The precision in the identification of these crops is highly dependent on the stage of the crop, the agroecological zone where most of the climatic characteristics are similar, as well as using one cropping season (Kumar *et al*., 2022), as was the case for this particular study. This can be explained by the reports of Vinciková *et al*. (2010), who indicated that the spectral characteristics of crops are influenced by the chlorophyll and water content and therefore change over the course of the growing season. The best results for crop identification may be achieved during the phase of full vegetation development, when the soil's influence on the habitat's spectral reflectance is lowest. However, the exact timing of the full development stages differs from crop to crop (Vinciková *et al*., 2011). To ensure an accurate result, the study used the vegetative stage of green gram and sorghum, where the crop was probably experiencing maximum growth and was adequate cover to reduce the influence of soil characteristics. Ultimately, crop area, crop identification, and yield data estimation would be accurate and easy to generate over large areas with correct spectral reflectance values. This research finding confirms the possibility of crop data estimation using earth observation data.

4.1.4 Landcover classification, crop area estimation and accuracy assessment

The data for cropping pattern data and spectral signature results obtained were used in the landcover classification as shown in Figure 4.24a and b. The cropping pattern data was used in training the classification as well as running the accuracy assessment of the final classification. Spectral signatures for green gram and sorghum were used in identifying areas that had crop during classification.

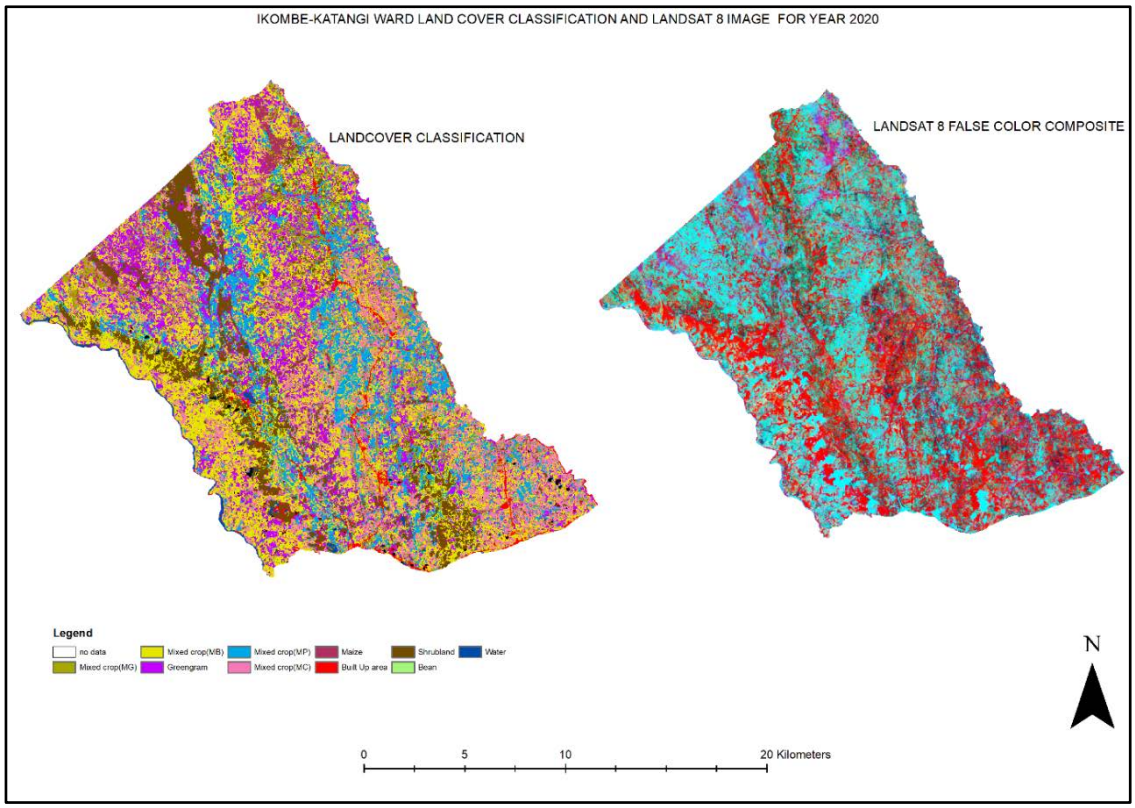


Figure 4.6a: Landcover classification of Ikombe-Katangi study area for the year 2020 OND rains season

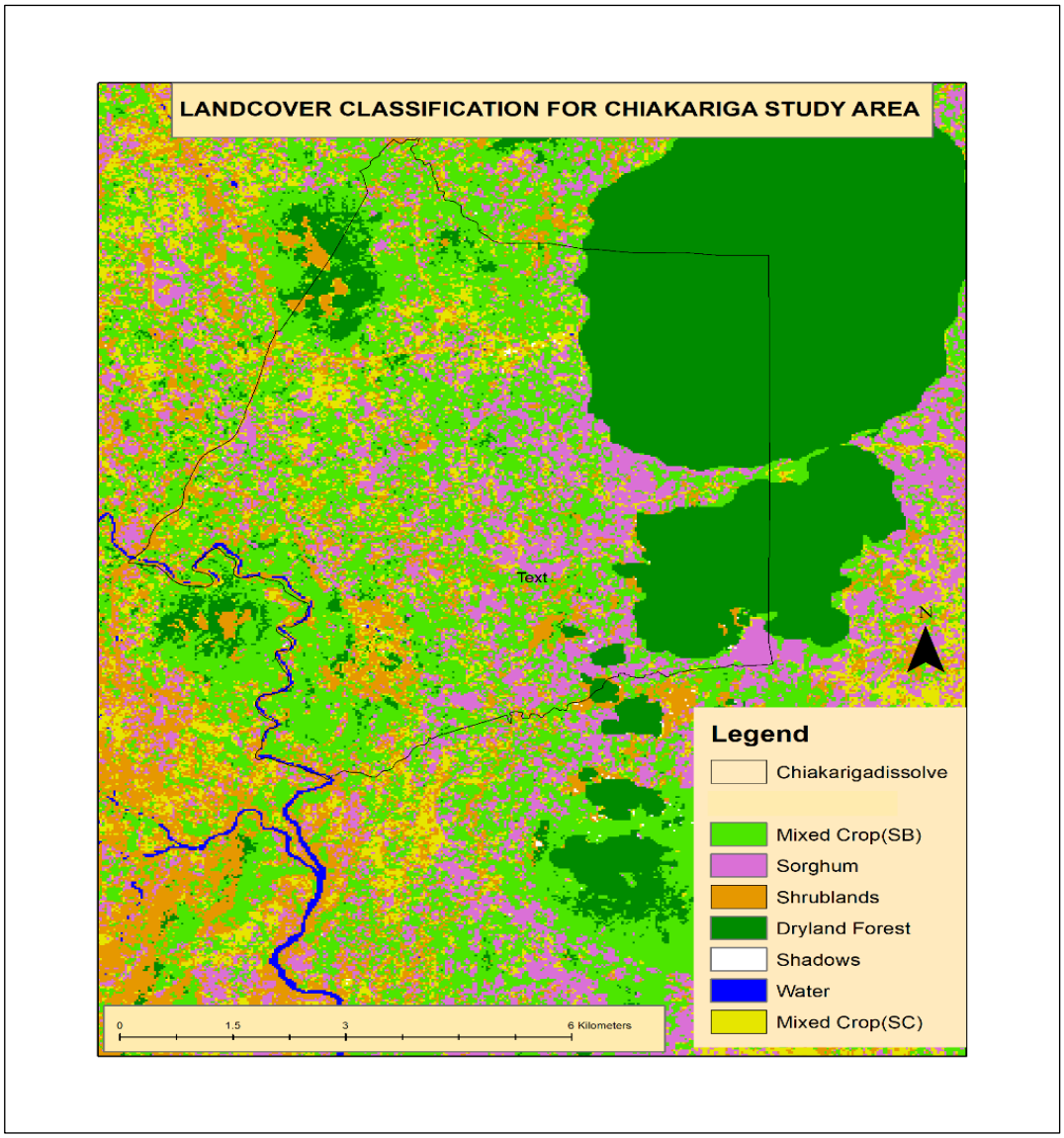


Figure 4.6b: Landcover classification of Chiakariga study area for the year 2020 OND rains season

A study by Forkuor *et al*. (2018) has significantly proven the importance of various bands in a land cover classification that features other vegetation alongside crops. Jamal-Eddine *et al*. (2018) showcased the good performance of pre-processed OLI-8 satellite data in the classification of very fragmented agricultural landscapes. The cropping pattern data generated was used in supervised classification using a maximum likelihood classifier to generate the cropped area. The results of the cropped

area are shown in Tables 4.5a and b. The crop area estimations in the Ikombe-Katangi area were 2445.93 ha, 10034.46 ha, 5981 ha, 4697.82 ha, 3743.82 ha, 586.35 ha, and 98.37 ha for mixed crop maize green gram (MG), mixed crop maize bean (MB), green gram, green gram (MG), mixed crop maize pigeon pea (MP), mixed crop maize cowpea (MC), maize, and beans, respectively, as shown in Figure 4.4a for the year 2020. The crop area estimations in the Chiakariga study were 1988.46 ha, 961 ha, and 469.62 ha for cropping patterns: mixed crop sorghum bean (SB), sorghum, and mixed crop sorghum, respectively, as shown in Tables 4.5a and 4.5b for the year 2020.

Table 4.5a: Area estimation for cropping pattern data in the Ikombe-Katangi study area

Class	**Distribution Pixel**	**Area in Ha**
Mixed crop(MG):	27,177	2445.93
Mixed crop(MB):	111,494	10034.46
Green gram:	66,461	5981.49
Mixed crop(MP):	52,198	4697.82
Mixed crop(MC):	41,598	3743.82
Maize:	6,515	586.35
Built Up Area	7,316	658.44
Shrubland:	36,780	3310.2
Bean:	1,093	98.37
Water:	3,041	273.69
No Data Area	1,304	117.36

Table 4.5b: Area estimation for cropping pattern data in the Ikombe-Katangi study area

Class	Pixel Distribution	Area in Ha
Mixed Crop(SB)	22094	1988.46
Sorghum:	10,685	961.65
Shrublands:	6,332	569.88
Dryland Forest	15394	1385.46
Shadows:	49	4.41
Water:	256	23.04
Mixed Crop(SC):	5218	469.62

Earth observation data have proven to be an effective tool for crop area estimation through crop classification (Castillejo-González *et al*., 2009). The contribution of crop area to crop yield estimation accuracy cannot be ignored (Gao and Zhang, 2021). The study incorporated crop area estimation as a step in the development of the crop yield tool for the Ikombe-Katangi and Chiakariga study areas. Maurya (2011) combined MODIS data and GIS techniques to undertake crop area estimation for soybean. This study concluded that the technique and data used enhanced the quantitative estimation of acreage and production of soybean. In another study by Kumar *et al*. (2022), microwave sentinel data was used in the acreage estimation of sugarcane, which increased the accuracy of the area estimation. The research incorporated cropping pattern data, which increased the accuracy level of data estimation. He study by Li *et al*. (2021) on yield estimation showcased that cultivar information in Irish potatoes was important for crop area estimation. This study confirms the importance of cropping data in yield and crop area estimation, which were used in this research to develop the crop model.

The accuracy assessment of the landcover classification of the study area was 83.3% with Kappa coefficient of 0.8066, Table4.6a while for sorghum was 82.58% with

Kappa Coefficient of 0.7845. Table 4.6b. The Producer accuracy and Users accuracy were both 82.86 and 85.29 respectively for the green gram crop while for sorghum it was 85.37% and 94.59% respectively.

Table 4.6a: Accuracy assessment result for landcover classification for IKombe-Katangi

Overall Accuracy	82.61%	
Kappa Coefficient	0.794	
Class	Producers Accuracy.	User Accuracy.
	(Percent)	(Percent)
Green Gram	75	100
IrrigatedCrop	60	100
Othercrops	100	56.25
OtherVegShrub	100	100
OtherVegTree	100	100
Towns	80	100
Water	100	100
Baresoils	33.33	100

Table 4.6b: Accuracy assessment result for landcover classification for Chiakariga

Overall Accuracy		82.6%
Kappa Coefficient		0.7845
Class	Producers Accuracy	Users Accuracy
	(Percent)	(Percent)
Sorghum	85.37	94.59
Mixed Crop(SB)	77.78	72.41
Shrublands	84.21	94.12
Dryland Fores	75	75
Mixed Crop(SC)	77.78	70
Water	93.33	100
Total	100	100

The overall accuracy was 84%, with producer and user accuracy of 82.26% and 85.29%, respectively, for the green-gram crop in the landcover classification, which confirms these approaches capabilities in crop identification. Other cropping patterns

whose identification was equivalently good were mixed crop (MP), mixed crop (MB), mixed crop (MC), and bean, whose producers and users accuracy were 79.31%, 88.89%, 70.59%, 83.33%, 100%, and 100%, respectively. The evaluation of other classes, such as maize, showed that 61% of producers' accuracy and 100% of users' accuracy This is an indication that the identification of maize could have been influenced by its presence in other cropping patterns. Tariq *et al*. (2022) observed that mixed pixels reduce the impact of mapping accuracy on land covers. Other non-crop classes were easily discriminated against, and their accuracy was good, as shown in Tables 4.6a and b. The classification approach differentiated other classes very well, where built-up area and shrubland classes were well discriminated with producers and users accuracy of 92.86% and 92.86%, 83.33%, 522, and 90.9%, respectively. Upon the classification of the Chiakariga study area, the following classes were identified: sorghum and mixed crop (SB). mixed crop (SC), water, forest, and shrublands, as shown in Figure 4.6b.

This multi-approach to landcover classification, especially in cropland, agreed with the work of Ahady and Kaplan (2021); Crnojevic *et al*. (2014); Mtibaa and Irie (2016); and Shi *et al*. (2020). It is a surefire way to make medium-resolution satellite imagery data better at identifying crops.This research builds on the need to develop spectral signature libraries in Africa, with more research on different satellite datasets. confirmation that green gram crop identification is a step towards the development of spectral signature libraries for Africa. Spectral signature libraries are designed for agricultural research purposes, such as the SPECCHIO Spectral online database maintained by the Remote Sensing Laboratories in the Department of Geography at the University of Zurich (Hueni *et al*. 2009).

4.2 Micro-agroecological zone data results for the study area

4.2.1 Analytical Hierarchy Process (AHP) results for green gram and sorghum

The AHP results in figures 4.7 a and b and 4.8 a and b for green gram and sorghum document the effect of crop performance in the study area. Based on these results, it was clear that surface temperatures and rainfall were the highest, i.e., 18% and 15.6%,

respectively, for green gram. Other parameters that scored the highest values were 12.3% and 13.9% for soil moisture and soil temperatures, as shown in Table 4.7a for green gram and 4.8 for sorghum. The rest of the parameters had scores ranging between 4.1% to 9.1% for green gram and sorghum, as shown in Tables 4.8a and b. The rankings are illustrated in Figures 4.7 a and b and 4.8 a and b for green gram and sorghum, respectively. To find the microclimatic zones in agro-ecological zones IV and V for the Ikombe-Katangi and Chiakariga study areas, the weighted scores were used in the weighted overlays. The MCDA has been widely used in different studies to analyze the suitability of crops for different agroecological zones as well as regions. Ayehu and Besufekad (2015) document the successes of weighted sum overlay in the identification of suitable areas for growing rice. The research points out that the MCDA approach using the AHP process could provide information to farmers that would help in the selection of cropping patterns locally. The current research is complemented by Victor and Samson (2019), who did a similar investigation for rice cultivation in Nigeria. The investigation in this study confirmed the successes of GIS-based MCDA support in suitability analysis for crops across regions. Using the GIS-based MCDA approach and AHP for identification of good land for different crops has worked well in a number of studies (Akinci *et al*., 2013; Mistri and Sengupta, 2020; Mugo *et al*., 2016; Tercan and Dereli, 2020; Ustaoglu *et al*., 2021). The overall conclusion from these studies is that the MCDA approach can help decision-makers accurately identify the requirements of selected crops. Once the ranking of parameters was achieved, a reclassification matrix was done, as shown in Tables 4.10 and 4.11 below. On parameters to use during the suitability analysis of a crop, Tercan and Dereli (2020) pointed out that slope, elevation, land use capability, and average temperatures were the main constraints in citrus production. Mugo *et al*. (2016) on the other hand used soil depth, elevation, slope, and aspect, which they deemed important in understanding suitable areas for rowing green grams in Kenya. In another study by Selim *et al*. (2018), the significance of soil parameters in avocado production is cited as critical. The lack of such datasets would result in spatial uncertainties. The soil parameters highlighted as critical in this research were soil pH, soil texture, soil

organic levels, and soil moistening conditions. A study by Yin *et al*. (2020), which focused on improving the weighted linear combination of MCDA and AHP, documents the need to include socioeconomic factors such as current use and regional differences. In our current study, this was improved through the consideration of analysis for specific agroecological zones that have already been classified for specific crops. A report by Manzi and Gweyi-Onyango (2020), which highlighted changes in climatic conditions and the need to include other important parameters, complemented our current research for further evaluation of any existing sub-zones. Other parameters of importance in this study, such as soil moisture and land surface temperatures, and have been documented by Rossato *et al*. (2017) research. The current research showed the significance of soil moisture through the correlation of precipitation and average water storage in the soil. Further, Faisal *et al*. (2020) documented the importance of the effect of drastic weather patterns on winter agricultural activity and the need for its consideration for future agricultural sustainability.

Table 4.7 Eigenvector weights for green gram sub-zonation criteria

	Criterion	Comment	Weights	+/-
1	Rainfall	Extreme importance	18.0%	11.3%
2	Land Surface Temperature	Extreme importance	15.6%	6.0%
3	Soil Moisture	Demonstrated importance	12.3%	4.2%
4	Soil Temperature	Demonstrated Importance	13.9%	3.9%
5	Soil PH	Strong Importance	5.6%	2.9%
6	Soil Drainage	Strong Importance	7.1%	3.7%
7	Soil Texture	Strong Importance	5.1%	1.9%
8	Slope	Moderate importance	4.1%	2.4%
9	Elevation	Moderate importance	9.1%	4.3%

Eigenvalue				Lambda:	10.996	MRE: 44.3%
Consistency Ratio	0.37	GCI:	0.27		CR:	7.5%

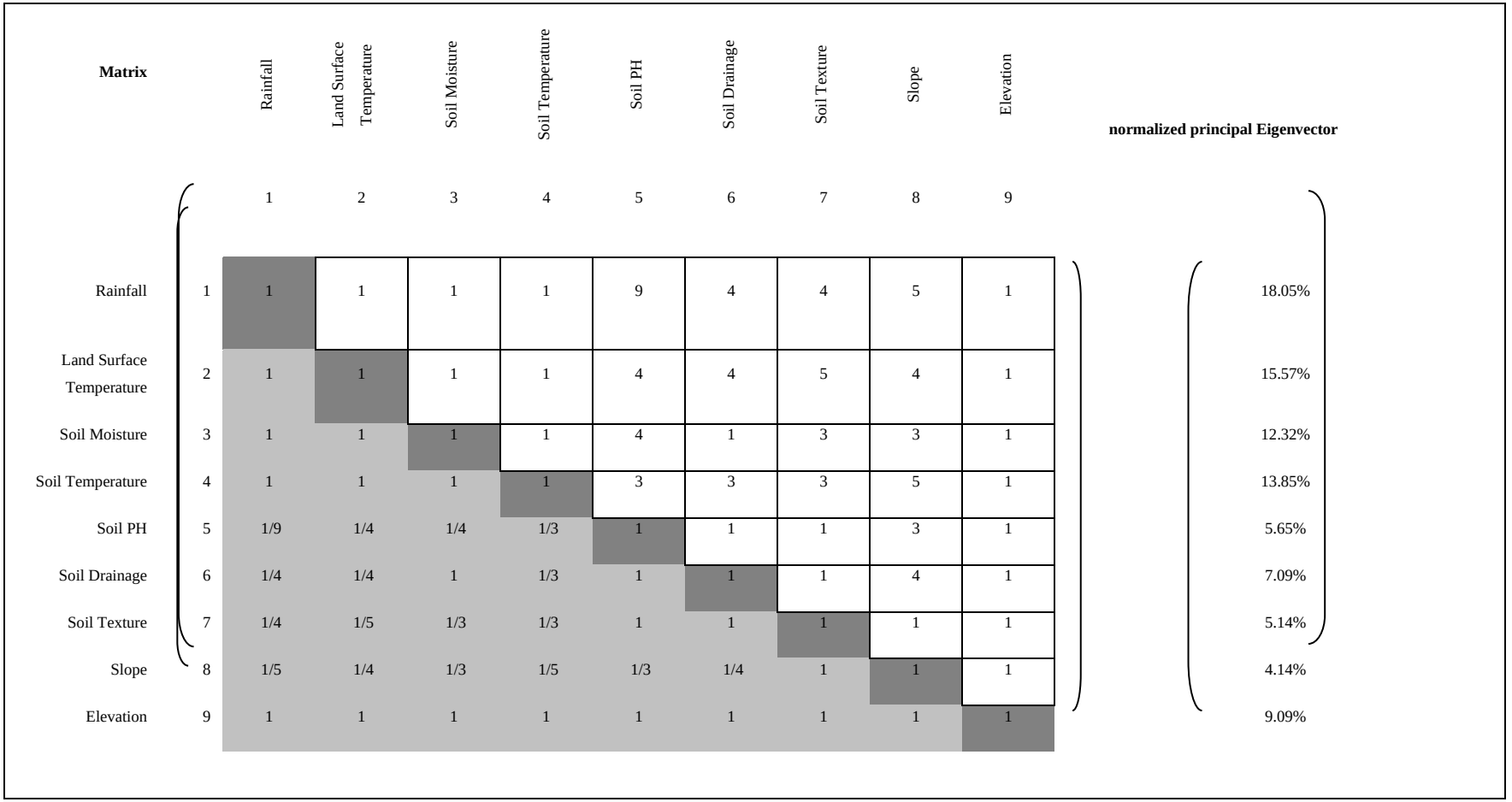


Figure 4.7a.0Pairwise comparison matrix for green gram sub-zonation criteria

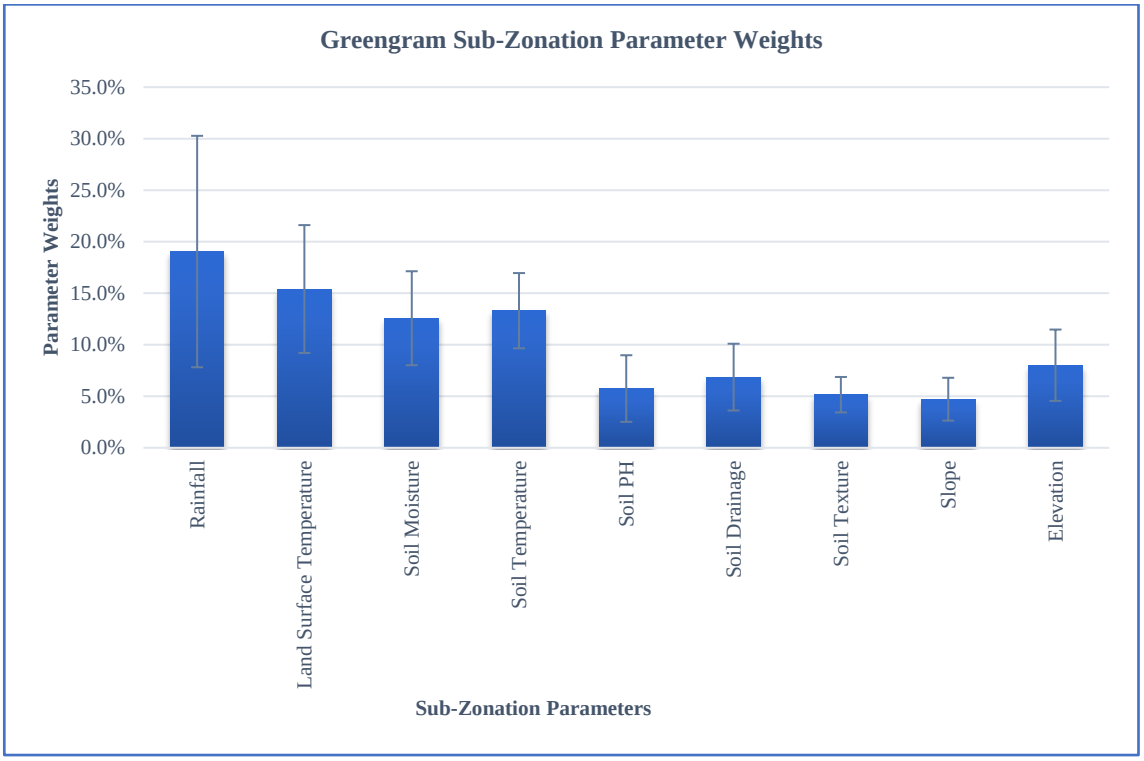


Figure 4.7b; Green gram sub-zonation parameter weights

The AHP process output for sorghum it is illustrated in Figure 4.8 a and b, which showcases the output weights. In sorghum parameters with the highest score were land surface temperatures and rainfall which were 19% and 15.4%. As indicated in Table 4.8, the other parameters that received the highest scores were soil temperatures and moisture, which came in at 13.9 and 12.3 %, respectively, and sorghum, which came in at 12.6 and 13.3 %, respectively, as indicated in Table 4.8 and Figure 4.8 aandb. The rest of the parameters had scores ranging between 4.7% to 8. % for sorghum was as shown on Table 4.8 which is illustrated in Figure 4.8a andb. The weighted scores were therefore used in the weighted overlays to identify the existing micro-climatic zones within agro-ecological zone IV and V for Chiakariga study area.The AHP process output for sorghum is illustrated in Figures 4.8 a and b, which showcase the output weights. In sorghum parameters with the highest score were, land surface temperatures and rainfall at 19% and 15.4%, respectively. Other parameters that scored the highest were 12.3% and 13.9% for soil moisture and soil temperatures,

as shown in Table 4.8, and 12.6% and 13.3% for sorghum, as shown in Table 4.8 and Figure 4.8 a and b. The rest of the parameters had scores ranging from 4.7% to 8.7% for sorghum, as shown in Table 4.8, which is illustrated in Figures 4.8a and b. The weighted scores were therefore used in the weighted overlays to identify the existing micro-climatic zones within agro-ecological zones IV and V for the Chiakariga study area.

Table 4.8 for eigenvector weights for sorghum sub-zonation criteria

	Criterion	Comment	Weights	+/-
1	Rainfall	Extreme importance	19.0%	11.2%
2	Land Surface Temperature	Extreme importance	15.4%	6.2%
3	Soil Moisture	Demonstrated importance	12.6%	4.6%
4	Soil Temperature	Demonstrated Importance	13.3%	3.7%
5	Soil PH	Strong Importance	5.8%	3.2%
6	Soil Drainage	Strong Importance	6.9%	3.2%
7	Soil Texture	Strong Importance	5.2%	1.7%
8	Slope	Moderate importance	4.7%	2.1%
9	Elevation	Moderate importance	8.0%	3.5%

Eigenvalue				Lambda:	10.895	MRE: 41.8%
Consistency Ratio	0.37	GCI:	0.24		CR:	6.7%

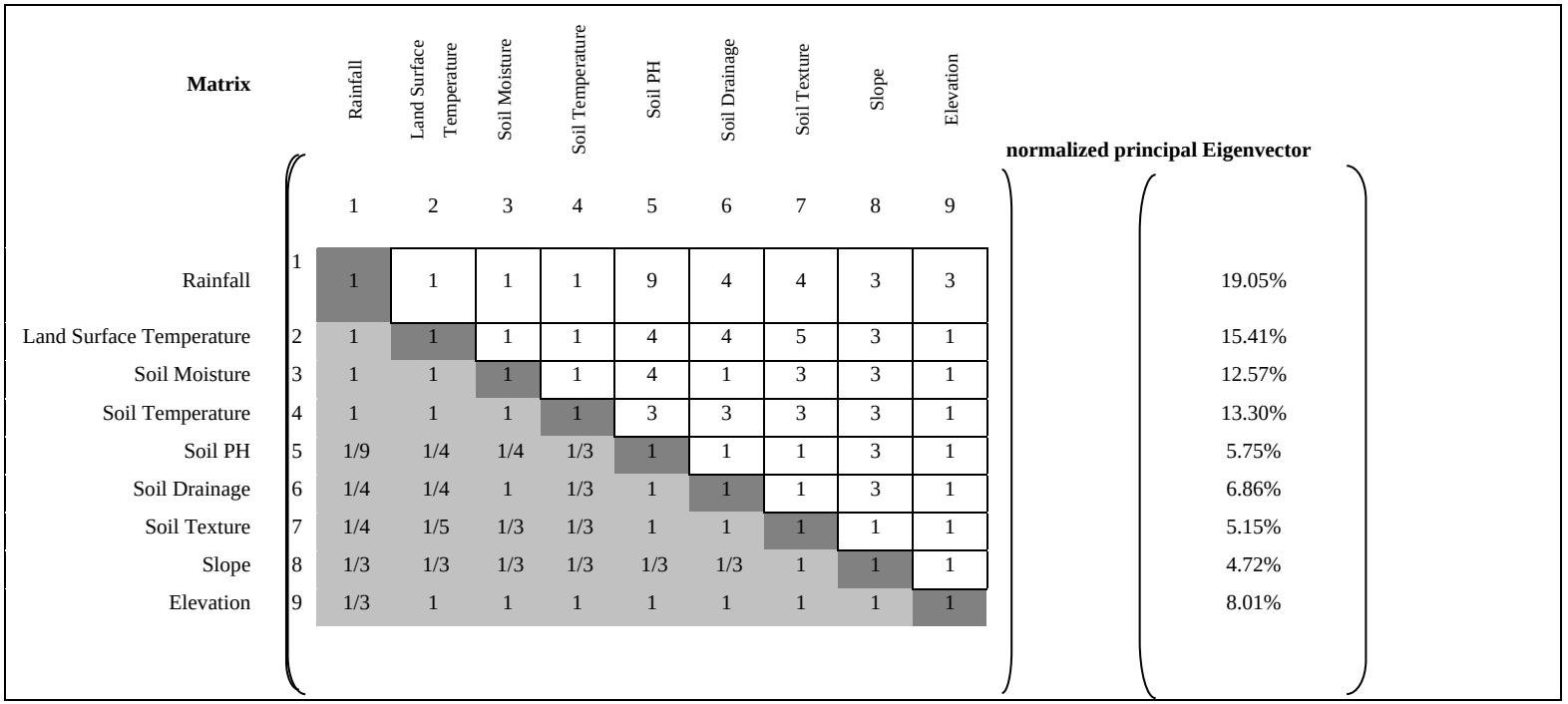


Figure 4.8a. Pairwise comparison matrix for sorghum sub-zonation criteria

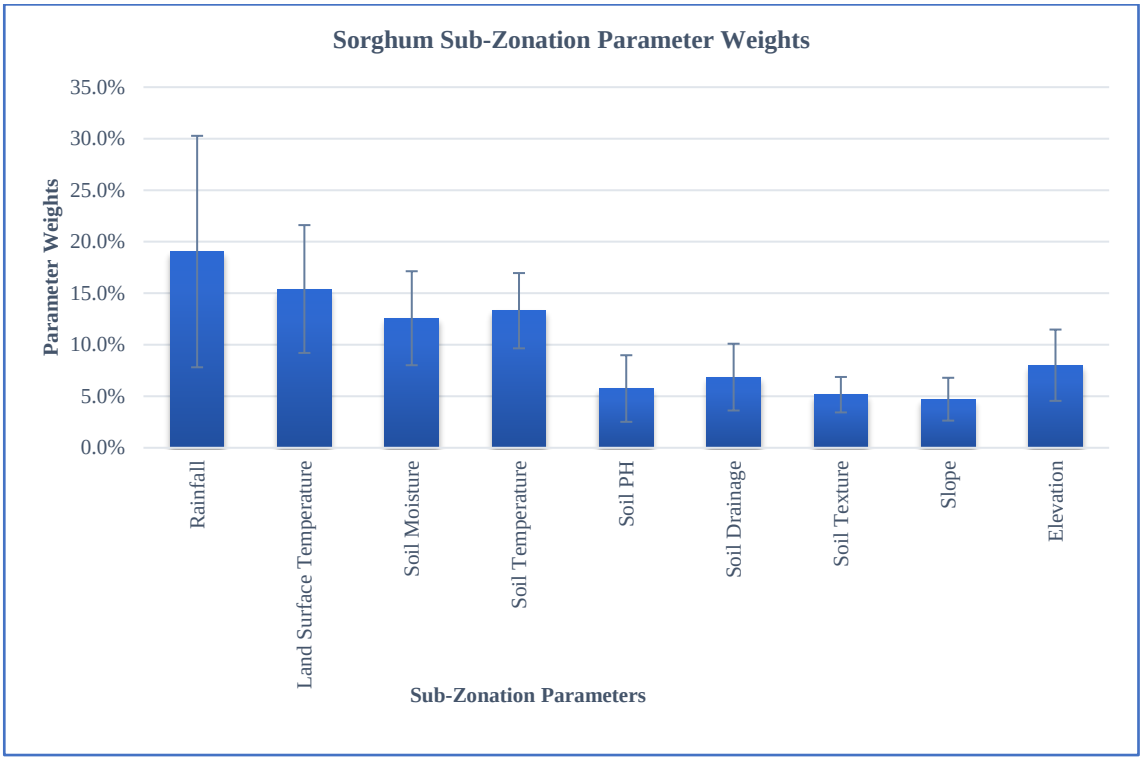


Figure 4.8b: sorghum sub-zonation weights

4.2.2 Reclassified and weighted Agroecological zone IV and V study area

The reclassified raster layers resulted in classes such as very poor, poor, moderately poor, good, moderately good, and very good, as visualized in Tables 4.10 and 4.11 for soil moisture, land surface temperatures, rainfall, soil temperature, soil drainage, soil texture, and soil pH in green gram and sorghum, with the overall subzone statistics shown in Table 4.9. The parameter raster layers used in the weighted overlay are shown in Figures 4.9 to 4.11 below.

Table 4.9: Agroecological zone lower midland IV and V subzone statistics

Zone	Soil PH	Soil Texture	Soil Drainage	Soil Temperature	Land Surface Temperatures	Rainfall	Soil Moisture	Elevation	Slope
1	5.0	1.0	5.0	23.0	22.0	1265.0	6.0	1500.0	3.4
2	6.0	1.0	5.0	23.0	30.0	1104.0	4.5	1318.0	4.1
3	7.0	6.0	5.0	23.0	36.0	618.0	5.7	461.0	5.3
4	7.0	6.0	5.0	23.0	36.0	527.0	5.6	423.0	5.5
5	6.0	6.0	5.0	23.0	36.0	527.0	5.6	432.0	5.5

Table 4.10: Reclassified parameters for green gram in agroecological zone IV and V

Classes Parameters	Class I(VP)	Class II (P)	Class III(MP)	Class IV (G)	Class V (MG)	Class VI (VG)
Land surface Temperature	<13^{o}c	13-23^{o}c	37-41^{o}c	34-37^{o}c	31-34^{o}c	23-31^{o}c
Soil Moisture	<2.0%	2.0-3.5%	>6.0%	5.5-6.0%	3.0-4.5%	4.5-5.0%
Soil Temperature	<5^{o}c	5-14^{o}c	14-20.5^{o}c	27-30^{o}c	25-27^{o}c	21-25^{o}c
Soil PH	<3.0	3.0-4.0	>7.0	4.5-5.0	5.0-6.0	6.0-7.0
Soil Texture	Sand, Clay	Silt clay	Silt clay loams	Silty loam, sandy loam, loamy sand	Clay loam, Loam	Sandy clay loams, sandy clay
Soil Drainage	Very poor	poor	Imperfect, moderately drained	Somewhat excessively drained	Excessively drained	Well drained
Rainfall	>1200mm	1000-1200mm	800-1000mm	600-800mm	200-400mm	400-600mm
Slope	>20%	16-20-%	12.0-16%	8.0-12%	4.0-8%	0-4%
Elevation	>1700m	1300-1700m	1000-1300m	700-1000m`	400-700m	<400m

Table 4.11: Reclassified parameters for Sorghum in agroecological zone IV and V

Classes Parameter	Class I(VP)	Class II (P)	Class III(MP)	Class IV (G)	Class V (MG)	Class VI (VG)
Land surface Temperature	<13ºc	13-23ºc	37-41ºc	34-37ºc	31-34ºc	23-31ºc
Soil Moisture	<2.0%	2.0-3.5%	>6.0%	5.5-6.0%	3.0-4.5%	4.5-5.0%
Soil Temperature	<5ºc	5-14ºc	14-20.5ºc	27-30ºc	25-27ºc	21-25ºc
Soil PH	<3.0	3.0-4.0	>7.0	4.5-5.0	5.0-6.0	6.0-6.7
Soil Texture	Sand ,Clay	Silt clay	Silt clay loams	Silty loam, sandy loam, loamy sand	Clay loam, Loam	Sandy clay loams, sandy clay
Soil Drainage	Very poor	poor	Imperfect, moderately drained	Somewhat excessively drained	Excessively drained	Well drained
Rainfall	>1200mm	1000-1200mm	800-1000mm	600-800mm	400-600mm	200-400mm
Slope	>20%	16-20-%	12.0-16%	8.0-12%	4.0-8%	0-4%
Elevation	>1700m	1300-1700m	1000-1300m	700-1000m`	400-700m	<400m

The spatial analysis of the Digital Elevation Model (DEM) was done and resulted in a derivative map of elevation as shown in Figure 4.9a. This was used as input to quantify the characteristics of the land surface, which is important in agroecological zone IV and V microclimatic zone analysis.

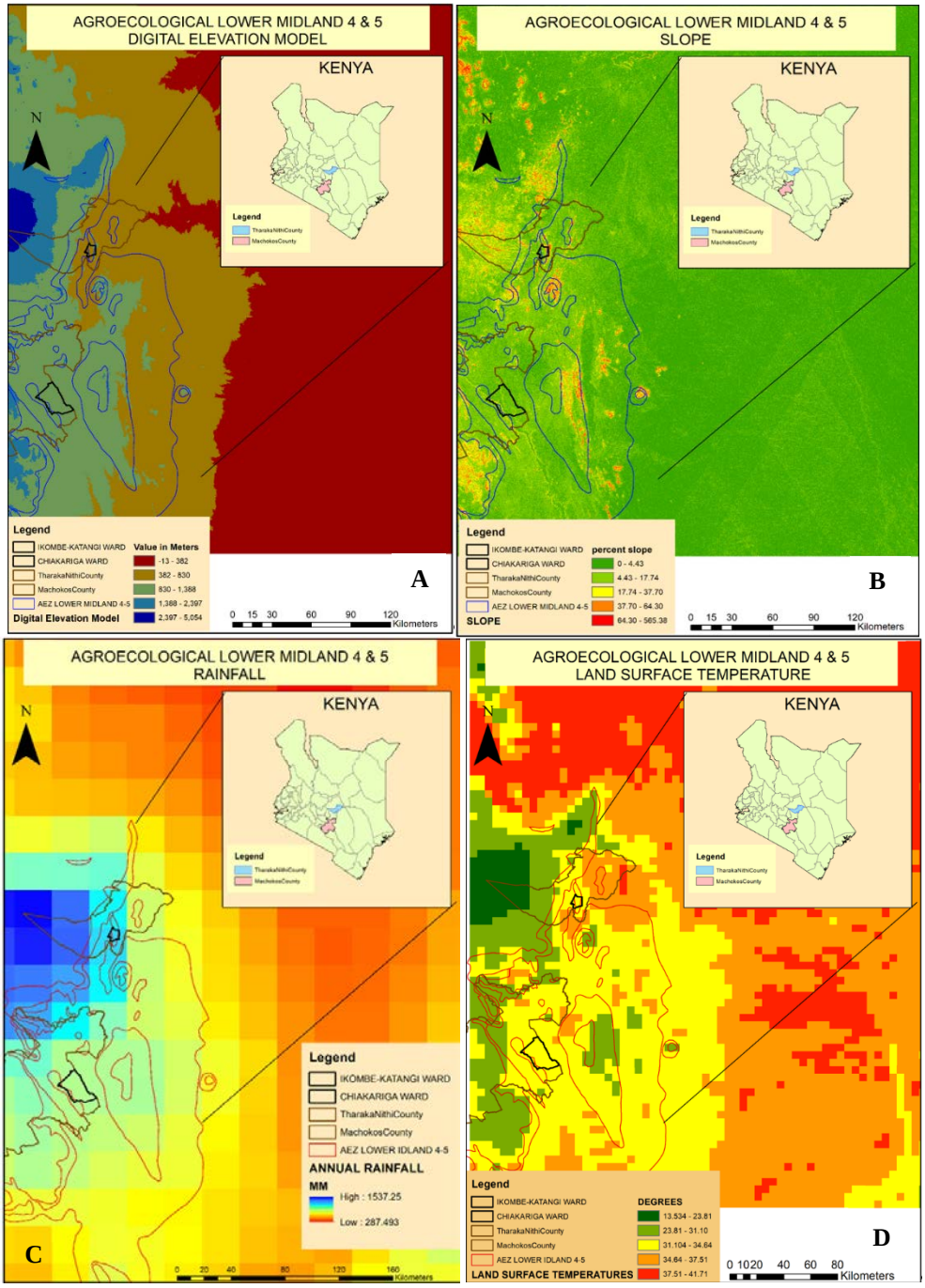


Figure 4.9: Agroecological zone IV and V DEM, slope, rainfall and land surface temperature ikombe-katanga and Chiakariga study area

A further analysis of the DEM-generated slope is shown in Figure 4.9 part B. The slope was calculated for each raster cell of the DEM using the elevation values of the cells and their neighbours. The elevations for the study areas were 830 to 1380 m and 382 to 832 m above the sea level for the Ikombe and Chiakariga study areas, respectively. The percent slope generated ranged from 4.43 to 17.74% for both the Ikombe and Chiakariga study areas, respectively. Generally, relatively lower slopes have been found to be better for cultivation than steeper ones (Fox *et al*., 1997). The slope characteristics were also classified based on the crop requirements for both sorghum and green gram. The slope requirement for the green gram, as documented by Mugo *et al*. (2020), shows that the crop performs well at 50–600 meters above sea level. On the other hand, based on the sorghum manual for semi-arid areas of Kenya by the Ministry of Agriculture, Livestock, and Fisheries Kenya, the crop grows well at an elevation of 1500 m or less above sea level. The slope most appropriate for green grams, according to Mugo *et al*. (2016), is 0–10%. In the sorghum-growing area, a suitable slope is not more than 10% for better crop production (Muzira *et al*., 2021). Fox *et al*. (1997) showed that steep slope gradients affect the infiltration of precipitated water; hence, it is very critical to consider slope when undertaking micro-agroecological sub zonation.

The rainfall parameter was also analysed in the ARCGIS environment and resulted in a derivative map for the study area, as shown in Figure 4.9-part c. The rainfall amount for Chiakariga was 1200 mm and above, while that of the Ikombe study area ranged from 600 mm to 800 mm. The impact of rainfall on crop production has been overemphasized by the various studies. For example, Kyei-Mensah *et al*. (2019) documented the impact of changes in intra-seasonal, inter-seasonal, and annual rainfall variability. Other studies, such as Mugo *et al*. (2016), retaliated on the significance of precipitation on crop productivity. Over the years, decreasing rainfall amounts in semiarid areas have led to a preference for sorghum as a crop (Muui *et al*., 2013; Dorcas *et al*., 2019). Abdullah-Al-Faisal *et al*. (2021) explored the impact of rising land surface temperatures on crop yield. The study documents that the changes in land surface temperatures have been closely associated with land use and land cover

changes over the years, with cropped areas showing significant changes over the years. The current study used rainfall and land surface temperature, which are very important as they influence crop growth. The land surface temperature was analysed, and the patterns for the study area were established as shown in Figure 4.9-part d. The land surface temperatures for the Chiakariga and Ikombe ranged from 31 to 34.6 ºC.

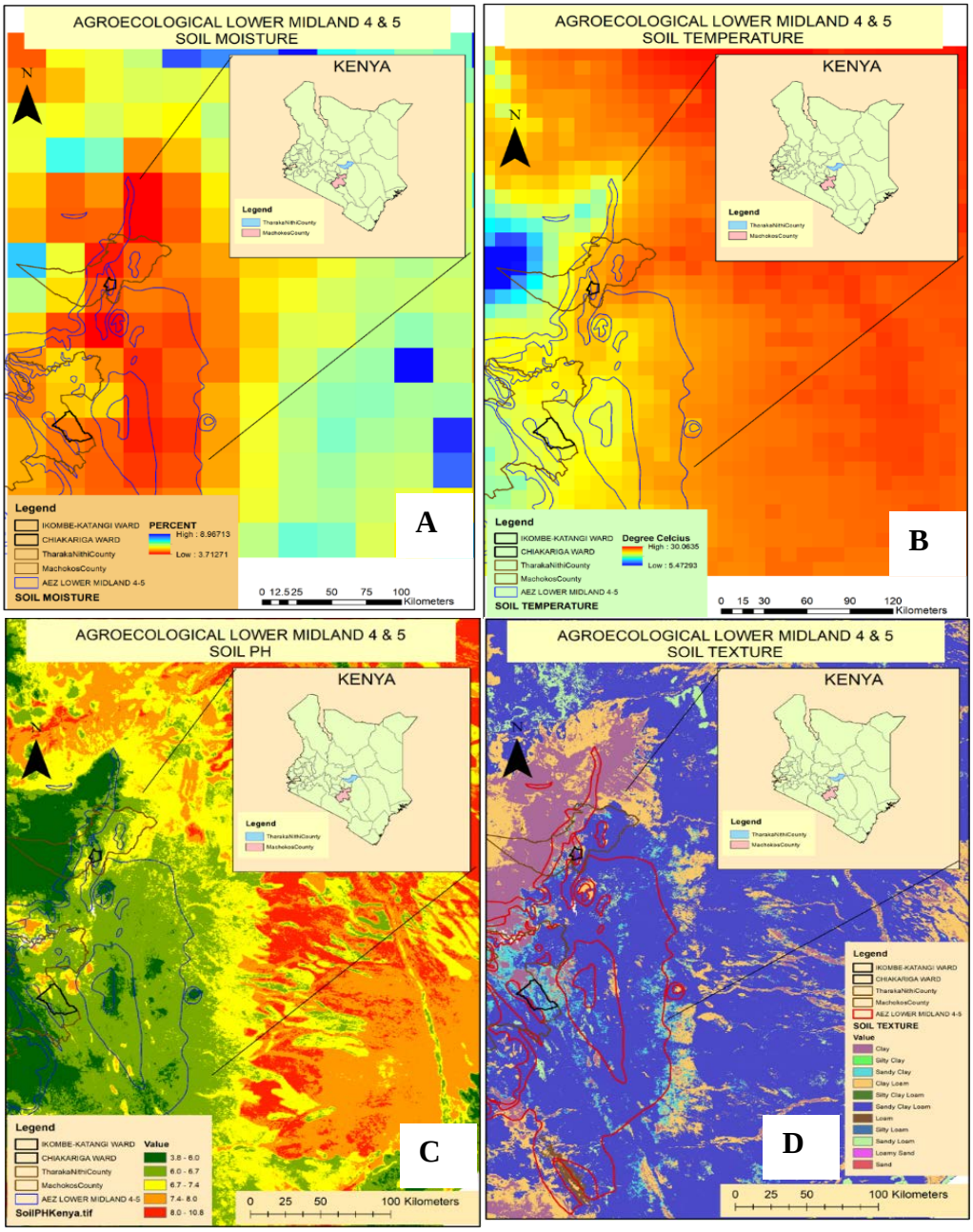


Figure 4.10: Agroecological zone IV and V results for soil moisture, temperature, pH and texture ikombe-katanga and Chiakariga study area

The soil moisture was extracted, and the results ranged from 3.7% to about 4.0% for the Ikombe-katangi and Chiakariga study areas, as shown in Figure 4.10 part a. The soil temperature was analyzed for the Ikombe-Katangi and Chiakariga study areas, as shown in Figure 4.10 part b. The results show soil temperatures ranging from 25 °C to 28 °C. The soil pH was assessed for the Ikombe-Katangi and Chiakariga study areas. The results showed that the soil pH ranged from 6.0 to 6.7 for both study areas, as shown in Figure 4.10 part c. The soil texture was assessed from the SOTWIS database, updated in 2021. The Ikombe-Katangi and Chiakariga study areas showed sand clay and sand clay loam textures, as shown in Figure 4.10 part d. The soil drainage was analyzed, and the results showed that the soils in the Ikombe-Katangi and Chiakariga study areas had well-drained soils, as shown in Figure 4.11.

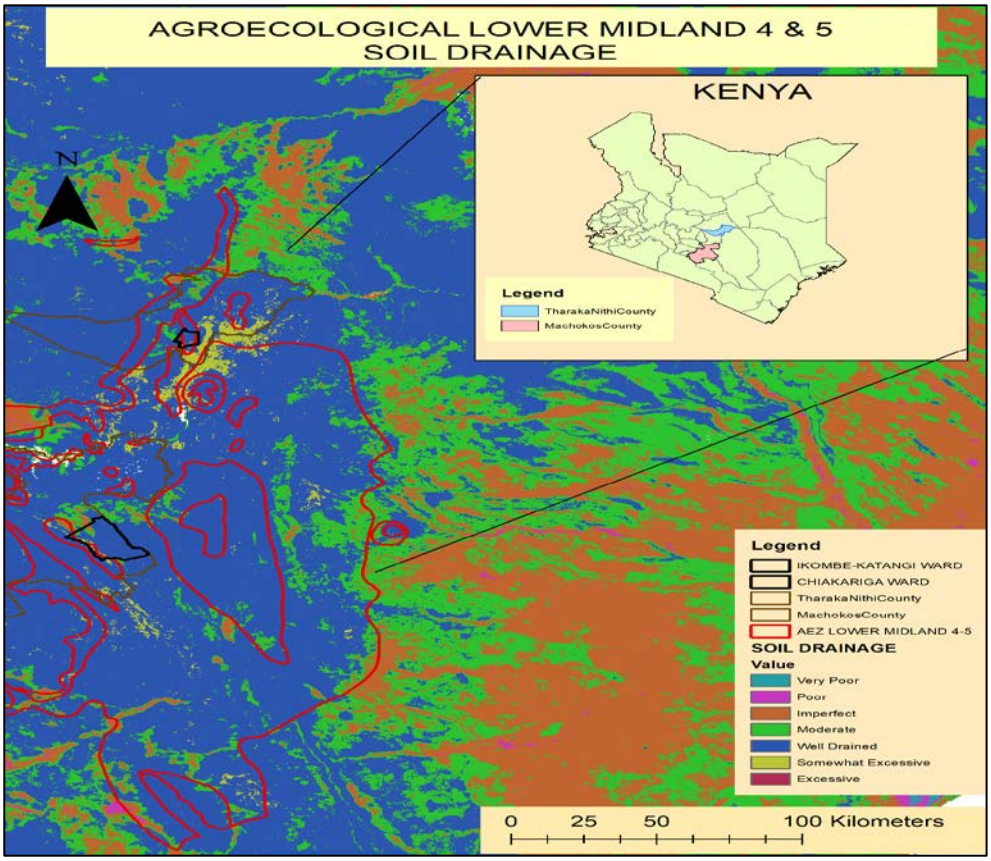


Figure 4.11: Agroecological zone IV and V results for soil drainage in ikombe-katanga and Chiakariga study area

The uptake of nutrients in sorghum is influenced by soil pH (Manzi *et al*., 2023a) and the drainage of the soils, as documented in the sorghum manual for semi-arid areas of Kenya by the Ministry of Agriculture, Livestock, and Fisheries Kenya. Checking these parameters is crucial for the performance of crops in the field. In addition, soil texture has been found to have a very profound effect on other soil properties, therefore influencing crop growth. Soil moisture and soil temperature influence the growth parameters of sorghum and green gram by providing the water and temperature levels required for appropriate crop growth (Manzi and Gweyi-Onyango, 2020). A study conducted in the Muooni catchment in the eastern parts of Kenya by Luwesi *et al*. (2017) showed declining soil moisture and drought severity, which have changed the ecological functions of the area. Bunn *et al*. (2015) did a multiclassification of the agroecological zone of Coffee arabica. Their research showed the increasing changes in different zones due to climate change. In addition, a study conducted in Egypt (Swelam *et al*., 2022) documented the effects of microclimates on irrigation requirements and water footprints in different regions. Further, the publication by Manzi and Gweyi-Onyango (2020) advocates for the use of additional parameters that were not considered in the initial zonation of agroecological zones, such as land surface temperatures, soil moisture, and soil temperature. In this research, the new parameters used are shown in figures 4.10 part c and d. In addition, topographical characteristics such as elevation and slope were also considered. Rainfall amounts and distribution, which have been experiencing drastic changes over the years due to climate change, were also put into consideration in developing the micro-climatic zones. Abd-elmabod *et al*. (2019) point out the need to use soil and climatic parameters alongside agricultural management databases for the purposes of land use planning to address issues of land vulnerability and suitability for farming. This research by Abd-elmabod *et al*. (2019) further points out how land degradation has affected soil productivity, which has ultimately had a negative impact on food security. In another study by Onduru and Preez (2008), parameters such as rainfall, soil pH, soil drainage, soil texture, and temperature were used in the investigation of suitable areas for crop irrigation. Wenner (1983) used, in the

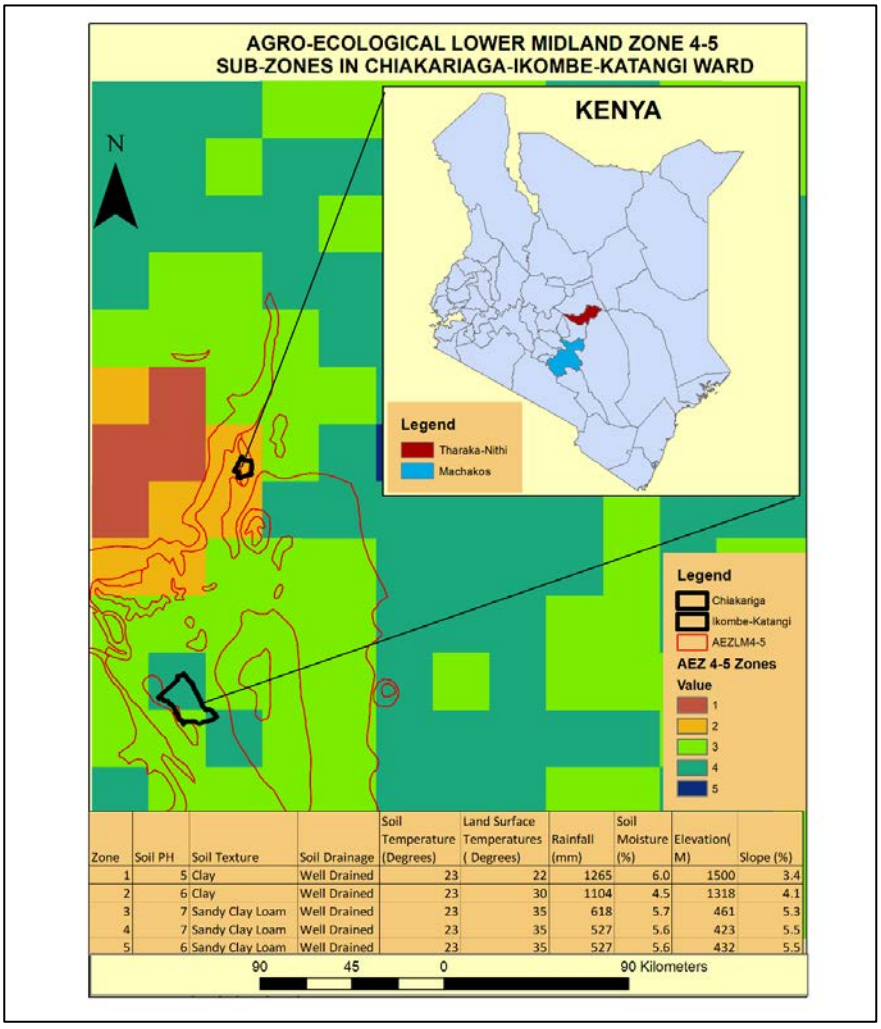
investigation of suitable areas of cropping in Nepal, soil texture, soil PH, soil drainage, and slope, among other factors. Studies by different workers (Mugo *et al*. 2016; Kahsay *et al*. 2018; Region 2018; Koomson *et al*. 2020) have reported and emphasized the importance of these parameters and the critical role they play in crop productivity.

The weighted overlay of the parameters analysed resulted in five subzones for agroecological zones IV and V in Tharaka Nithi and Machakos counties, as displayed in Figure 4.12, indicating the presence of micro-climatic zones in the study area. The parameters for each of the subzones were generated using spatial analysis and zonal statistics in the ArcGIS environment, as shown in Table 4.9. This research was able to define four subzones that exist within the agroecological zone lower midland IV and V of Tharaka Nithi and Machakos counties based on the parameters that were analysed, as shown in Figure 4.12. Sub-zone 5 fell outside the boundaries of lower midland zones IV and V, as shown on the map. The results of the investigation are displayed in table 4.12. According to the results, there was variation in land surface temperatures, where sub-zones 1 to 4 had mean land surface temperatures of 22^{o}C and 30^{o}C, while the rest had 36^{o}C, respectively, as estimated from the mean distribution of the area of study using ArcGIS zonal statistics. Land Surface temperatures, as described by Abdullah-Al-Faisal *et al*. (2021), influence photosynthesis, water and nutrient absorption, transpiration, respiration, and enzyme activity. Mugo *et al*. (2020) highlight that a suitable temperature for seed germination and proper growth in the case of green gram is between 28^{o}C and 30^{o}C. This is also the case with sorghum, which can survive in a wide temperature range, from 15^{o}C to 35^{o}C. Furthermore, the research confirmed that temperatures falling outside of this range are likely to negatively influence the growth of the crop. In the study by Mugo *et al*. (2020), most of Kenya, including the study area of the current research, were found to be moderately suitable for green gram production. This study carried out the sub-zonation of agroecological lower midland zones IV and V. It was evident that within the moderately suitable areas of green gram production, some pockets were established to be highly suitable. From Figure 4.9part d, based on the land surface

temperatures, one can identify that zones 3 and 4 were found to be more ideal for green gram production, while zones 2, 3, and 4 were suitable for sorghum production.

Mugo *et al*. (2020) research indicates that the optimum rainfall amount for green gram should be within 250–350 mm of rainfall, while for sorghum, it should be 250–400 mm, according to the "Enhancing Sorghum Semi-Arid Kenya" manual by the Ministry of Agriculture, Livestock, and Fisheries (MOALF). In addition, the moderately suitable areas had rainfall amounts of 350–600 mm (Mugo *et al*. 2020). Zones 3 and 4, whose land surface temperatures were suitable for the growth of green gram, had mean rainfall amounts of 618 and 527 mm, respectively. This rainfall range is also adequate for green gram production. In sorghum production rainfall amounts of about 600mm were adequate for germination and crop growth (Kahsay *et al*., 2018). A study conducted by Hossen *et al*. (2021) for suitable areas for mung bean production in Bangladesh and highlighted that elevations of 0–1600 m above sea level, optimum temperatures of 28–30ºC, and comparatively low rainfall amounts were favourable. Soil temperature levels of 20ºC to 25ºC have been found to be optimal for legumes, while 28ºC to 30ºC have been found to be adequate for cereals (Feyisa *et al*., 2020). Feyisa *et al*. (2020) showed that temperatures of 20ºC to 35ºC were optimum for the germination of the green gram crop, while the optimum seed germination and growth of sorghum were found to be 20ºC to 30ºC by Anda and Pinter (1994). In terms of soil temperature, all zones, i.e., 1 to 4, were found to be within a suitable range for optimum germination and growth of the green gram crop. This means that the parameter did not influence the sub-zonation in the MCDA analysis. This was also true for soil drainage, which was well drained in all zones. Soil moisture levels have been found to play a critical role in the germination of the seed and are highly dependent on the amount of precipitation as well as the precipitation rate, type of soil, and slope of the land. Rainfall amounts do not directly translate to the available soil moisture for crop growth (Li *et al*., 2014). The soil parameters used in this study, i.e., soil pH, soil texture, and soil drainage, were found to be good for the germination of green gram and sorghum crops. The soil pH for agroecological lower midland zones IV and V ranged from 6.0 to 7.0, and the soil

texture was sandy clay loams with a few pockets of clay soils (Figure 4.12). In addition, the soils were well drained. Figure 4.36 shows that the elevation ranges from 423 to 1500 m above sea level, while the slope is generally 3.4-5.5%, which is ideal for green gram and sorghum production. (Mashao and Prinsloo, 1994) indicate that the ideal soil characteristics for sorghum production are soils whose clay content is 10 to 30%, with a pH of 5.5, 50, or 8.0 being well tolerated. Further, the study points out that sorghum can tolerate a short period of waterlogging; otherwise, well-drained soils are good for sorghum.



Zone	Soil PH	Soil Texture	Soil Drainage	Soil Temperature (Degrees)	Land Surface Temperatures (Degrees)	Rainfall (mm)	Soil Moisture (%)	Elevation(M)	Slope (%)
1	5	Clay	Well Drained	23	22	1265	6.0	1500	3.4
2	6	Clay	Well Drained	23	30	1104	4.5	1318	4.1
3	7	Sandy Clay Loam	Well Drained	23	35	618	5.7	461	5.3
4	7	Sandy Clay Loam	Well Drained	23	35	527	5.6	423	5.5
5	6	Sandy Clay Loam	Well Drained	23	35	527	5.6	432	5.5

Figure 4.12: Results of Micro-Zone of agroecological zone IV and V of the study area

On the other hand, Karienye *et al*. (2019) point out that semi-arid and arid areas are good for sorghum production. The ecological requirements of green gram, according to Mugo *et al*. (2016), terms of soils should be well-drained loamy to sandy loamy soil, and the pH should be 6.0 to 7.0 in well-drained soil. Mugo *et al*. (2016) found that in the research for the identification of suitable land for green gram production in Kenya, optimum conditions were found to be well-drained soils with a pH of 6.2 to 7.0, loam to sandy loam soil, and a slope of between 0 and 10%. These studies confirm that agroecological Lower Midland Zone IV and V soil characteristics are good for sorghum and green gram production (Manzi and Gweyi-Onyango 2020). These research findings focus on the rainfall, land surface temperature, soil moisture, and soil temperature that have had an impact on the production of green gram and sorghum in the study area. It is confirmed that while some areas or zones remain good for the production of green gram and sorghum, other zones have changed due to the changing rainfall amount and land surface temperature over the years.

4.3 Sorghum and green grams crop yield results under farm field conditions

Ten parameters from remote sensing-based data were used in the estimation of sorghum and green gram yield under farm field conditions. The data generated for the various parameters is shown in Figures 4.13–4.19. The generation of the parameters used in yield estimation was also done and validated with field data collected as shown in Table 4.12. The field validation data is shown in Tables 4.15–4.20, while the validation results are displayed in Figures 4.20–4.24. The leaf area index was calculated using the EVI results and validated using field results from the two study areas. These results of the extracted parameters are shown in Figure 4.13a, b, and c, and 4.14a, b, and c for sorghum and green gram crops, respectively. The LAI ranged from 1.068 to 1.77 for green gram and 3.0 to 4.48 for the sorghum crop. The EVI results were 0.36 to 0.6 and 0.5 to 0.6, while the biomass ranged between 740 and 1490 kg/acre and 2160 and 1650 kg/acre for sorghum and green gram. Respectively. The average biomass was 1170 kg/acre and 1890 kg/acre for sorghum and green gram respectively. Leaf Area Index (LAI), which is the total green leaf area (double-sided) per unit horizontal ground surface area of vegetation canopy Watson (1958; Fernandes

et al., 2014; Manzi *et al*., 2023b) is an essential biophysical variable. It is used in soil-vegetation-atmosphere modeling (Anav *et al*., 2013; FAO, 2008; Launay and Guerif, 2005; Laurent *et al*., 2014).

Table 4.12: Field data results for the study area

CROP	SOIL PH	N in g/kg	P g/kg	K g/kg	YIELD IN KG	size of farm
Green grams	6.5	0.121	0.478	131.4	40	
Green grams	7.1	0.03	1.308	124.8	90	1
Green grams	7.8	0.029	0.465	116	5	1
Green grams	7.2	0.092	0.778	118.3	70	1
Green grams	7.5	0.096	0.098	125.8	10	1
Green grams	7.0	0.015	4.637	114.2	8	1
Green grams	7.9	0.028	1.970	136.4	40	1
Green grams	7.8	0.037	3.488	106.6	120	1
Green grams	7.8	0.035	2.049	128	25	1
Sorghum	7.8	0.006	0.305	47.5	100	1
Sorghum	7.6	0.052	3.125	107.9	450	1
Sorghum	7.9	0.02	0.872	123.8	1950	2

In agroecosystems, the total leaf area of the crop canopy is one of the key constraints on carbon assimilation and transpiration rates, which drive the accumulation of crop primary productivity (Nguy-Robertson *et al*., 2013). Therefore, LAI is required to estimate photosynthesis, evapotranspiration, crop yield, and many other physiological processes (Cao *et al*., 2015). The EVI has frequently been used for crop growth and yield-related research as a remote sensing parameter (Huang *et al*., 2019; Kogan *et al*., 2012; Liu *et al*., 2014; Zhang *et al*., 2014). EVI-based crop growth metrics are much closer to capturing crop status and growth characteristics, and growth metrics

can be much more correlated to crop yield than an NDVI (Huang *et al*., 2019). Similarly, this study used EVI indices to ensure that the crop yield estimation would be accurate.

4.3.1 EVI, LAI and biomass results for green gram and sorghum

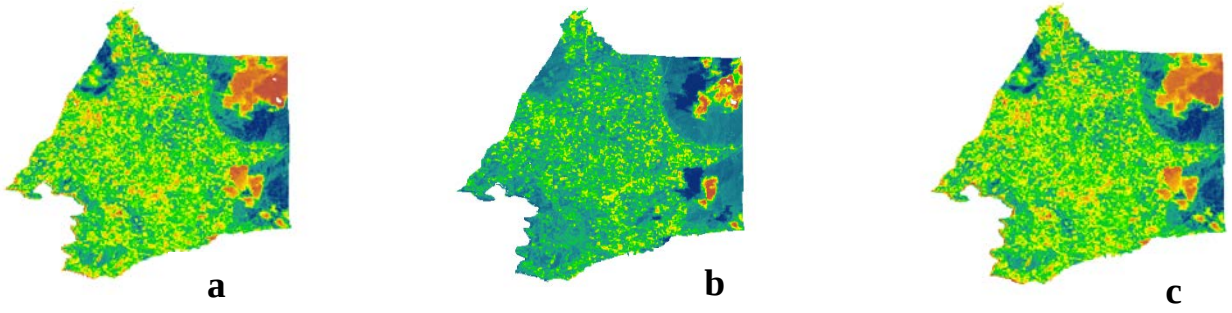


Figure 4.13a, b andc: Leaf Area Index, Enhanced Vegetation Index and Biomass results for Sorghum crop in Chiakariga study area

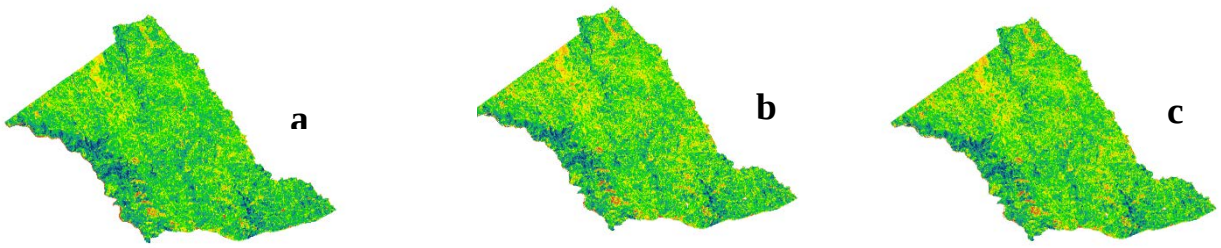


Figure 4.14a, b andc: Leaf Area Index, Enhanced Vegetation Index and Biomass results for Green gram crop in Ikombe-Katangi study area

LAI is one of the important parameters that reflect the crop growth stages of vegetation and is widely used in the quantitative analysis of crop models (Parker, 2020; Yan *et al*., 2019). (Parker, 2020) reiterates on the importance of leaf area index (LAI) by discussing remote sensing models used in its estimation. Several studies have extracted biophysical parameters from satellite imagery (Hui and Yao, 2018; Yu *et al*., 2019) and assimilated them into simulation models. Kang *et al*. (2016)

successfully assimilated Landsat-derived LAI time series into crop model simulations using ensemble Kalman filters for individual fields or pixels. The EVI ranges for green gram and sorghum were found to be 0.36–6.0 and 0.5–6.0, respectively, which is typical of many crops at that vegetative stage (Shammi and Meng, 2021). The biomass values obtained were found to be within the normal range of wet biomass at the vegetative stage in semi-arid areas for sorghum and green gram (Service and Moines, 2018). The covariance correlation matrix between the extracted and observed data for the LAI was R^2 = 0.93 and R2 = 0.99 for green gram and sorghum, respectively, as shown in Figure 4.15a and b. The results used in the correlation matrix are shown in Table 4.13.

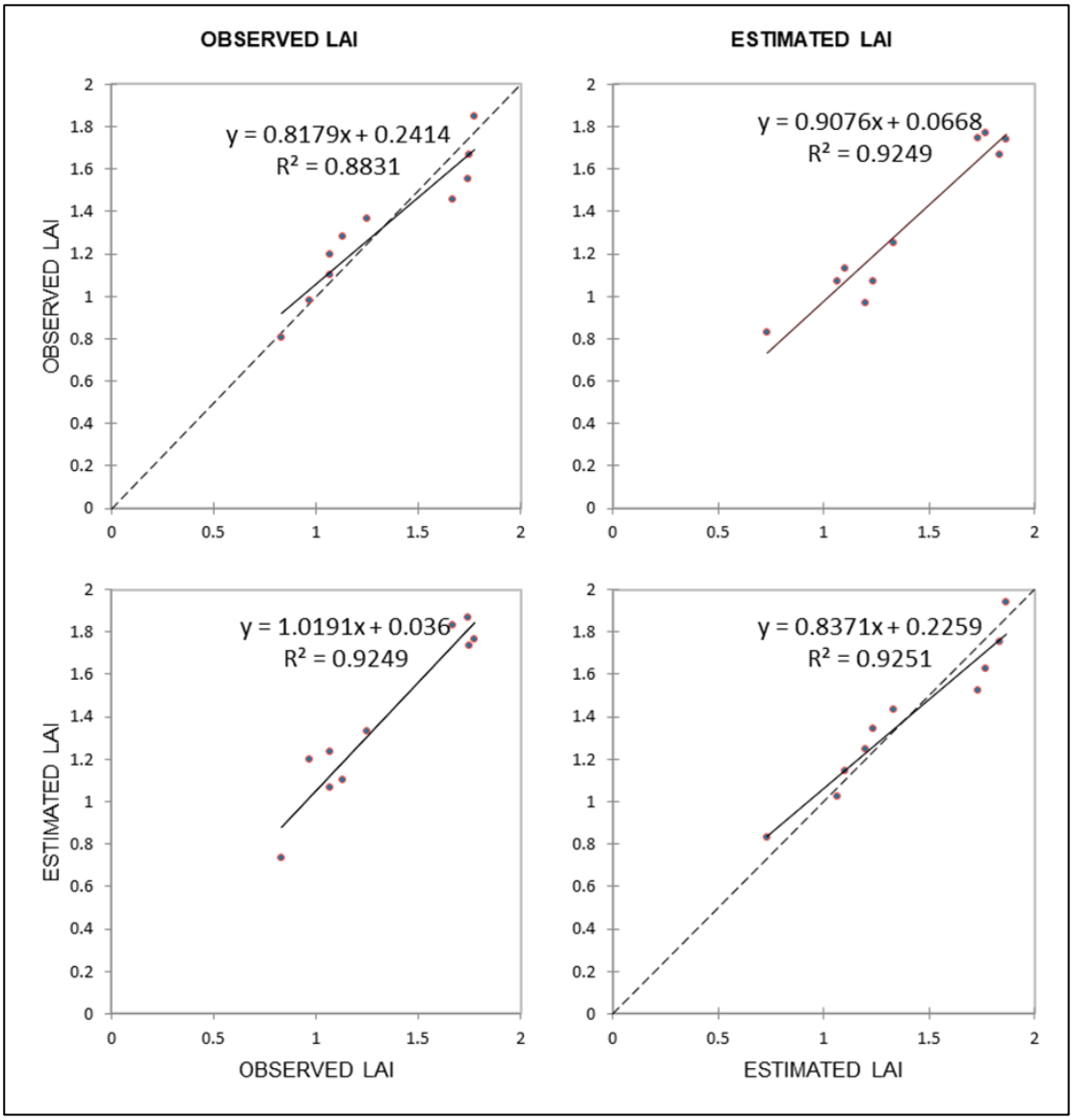


Figure 4.15a: Regression results for remote sensing-based leaf area index correlations with green gram field data

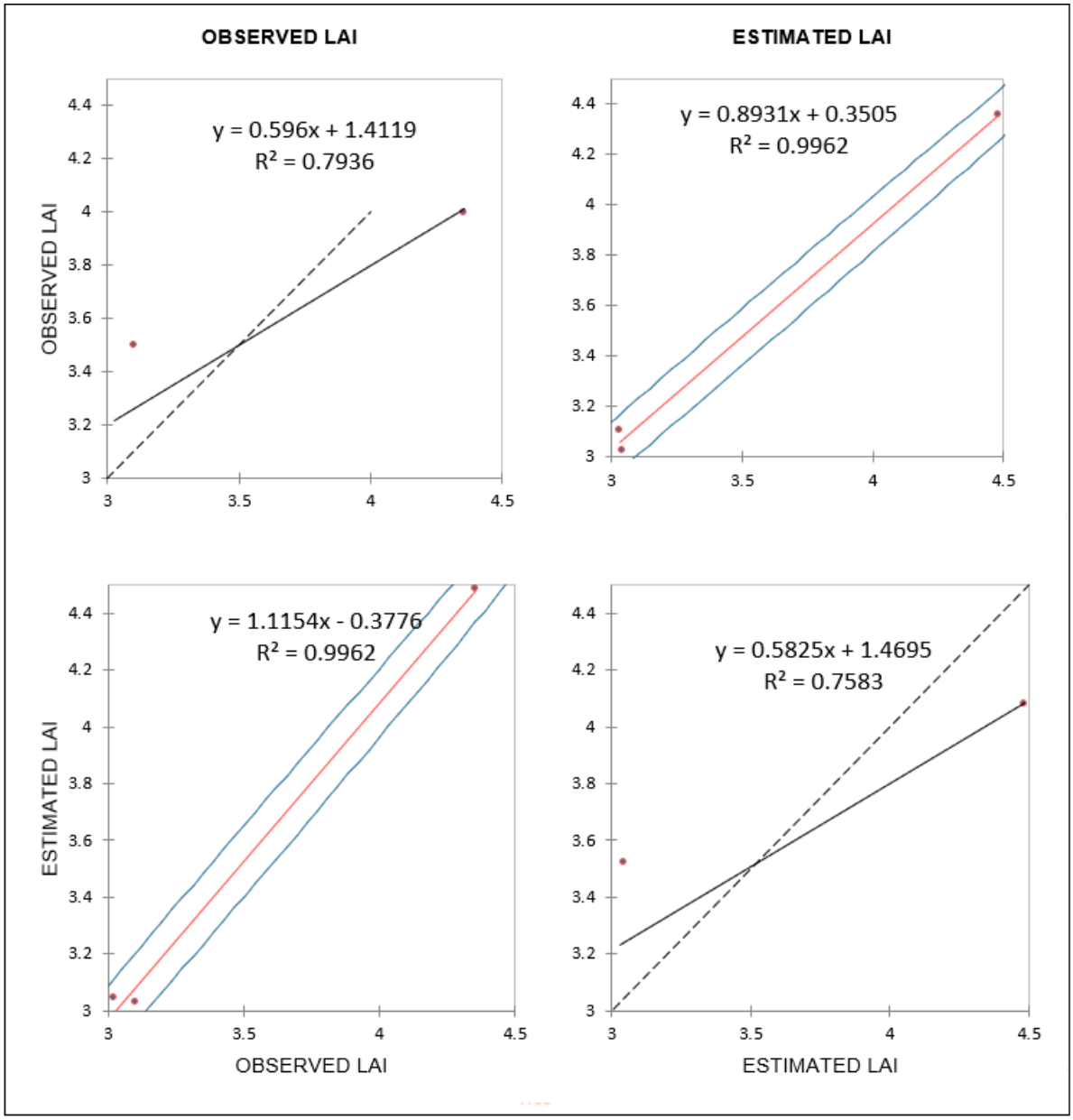


Figure 4.15b: Regression results for remote sensing-based leaf area index correlations with sorghum field data

Table 4.13: Field and remote sensing extracted Leaf Area Index Data

Crop	Observed LAI	Estimated LAI
Green Gram	1.068	1.2333
Green Gram	1.13315	1.100
Green Gram	1.7425	1.8667
Green Gram	1.07165	1.0667
Green Gram	1.25095	1.3333
Green Gram	1.67	1.8333
Green Gram	1.77275	1.7667
Green Gram	1.7472	1.7333
Sorghum	3.1023	3.02972
Sorghum	3.02451	3.04582
Sorghum	4.3567	4.48495

4.3.2 Soil organic carbon, Land surface temperatures and soil moisture results for green gram and sorghum study areas

Soil organic carbon, land surface temperatures, and perpendicular soil moisture index were extracted as shown in Figures 4.17 a, b, and c and 4.18 a, b, and c for sorghum and green gram, respectively. The extracted values for soil organic carbon ranged from 0.1369–16.0325 g/kg and 10.991–18.377 g/kg for sorghum and green gram, respectively. The values for land surface temperature that are closely inversely correlated to the soil moisture index ranged from 25.37–32.69 ^{o}C and 25.2-26.52 ^{o}C for sorghum and green gram crops, respectively. The average land surface temperatures for sorghum and green gram were 26 ^{o}C and 28.69 ^{o}C, respectively. For the soil moisture index, the values were 0.516–1.183% and 0.48–0.586% for sorghum and green gram, respectively. Soil organic carbon plays a role in enhancing soil biodiversity by enabling essential microorganisms to survive and interact with one another (Orgiazzi *et al.*, 2016). Further, soil biodiversity is responsible for the formation of soil organic matter that contributes to soil organic content (FAO and ITPS, 2015). The soil organic content determines the number and activity of soil biota

responsible for the microbial community that interacts with plant roots (Thiele-Bruhn *et al*., 2012). The levels of organic soil in this study had the implication of decreasing soil fertility and this is supported by results of Pichot *et al*. (1981), whose study in the semi-arid areas of Burkina Faso showed decreasing soil organic carbon levels against high levels that would enhance crop yield. Pichot *et al*. (1998) concluded that different soil fertility regimes during farming affected the expected soil organic carbon levels. Recent developments in geostatistics, artificial neural networks, and multiple regression have made it possible to accurately estimate the spatial variability of SOC through digital soil mapping (Mishra *et al*., 2008). The digital soil maps can be done at plot level (Simbahan *et al*., 2006), watershed level (Ng *et al*., 2019), and regional scales, where the potential for SOC can be analyzed (Minasny *et al*., 2013). The current study considered soil organic carbon due to its significant role in ensuring soil fertility. The significance of land surface temperatures in climate monitoring cannot be ruled out in extreme weather event scenarios, as reported by Sobrino and Jiménez-Muñoz (2005). Furthermore, Vlassova *et al*. (2014) emphasize land surface temperature as an important parameter in soil-vegetation transfer modeling in most terrestrial environments. This parameter becomes even more important in evaluating cropping seasons and assessing the length of growing periods. The land surface temperatures were within the expected range of the semi-arid regions of the current study area. The average soil moisture index for sorghum and green gram were 0.783% and 0.549%, respectively. The levels are typical of arid and semi-arid areas (Mohamed *et al*., 2020). The results from the study indicated low moisture levels, as would be expected from the expected averages. Further, Mohamed *et al*. (2020) indicate that the determination of soil moisture content is an effective factor in biological processes and the evolution of the soil profile. In addition, soil moisture affects vegetation distribution. Therefore, the lack of moisture may lead to drought and degradation, especially in rain-dependent areas. This soil moisture index was key in crop yield modeling since these areas are prone to drought and soil moisture determines crop growth at various stages. Kamara and Jackson (1997) state that soil moisture is a better indicator of water availability to the crop than rainfall amount. Improved soil moisture

management can only take place when analysis of soil moisture content is carried out and is well understood. Understanding soil moisture content scenarios assists in working toward high water retention in the soil, resulting in an eventual reduction in soil runoff in some cases (Kamara and Jackson, 1997). In semi-arid regions, the issue of soil moisture becomes very important since moisture stress in crops is one of the key determinants of crop failure. In other studies, Pellarin *et al*. (2020) have shown that soil moisture content can be used to infer precipitation using the PrISM model. This kind of modeling is believed to provide useful information concerning crop yield estimates and irrigation demands over large areas.

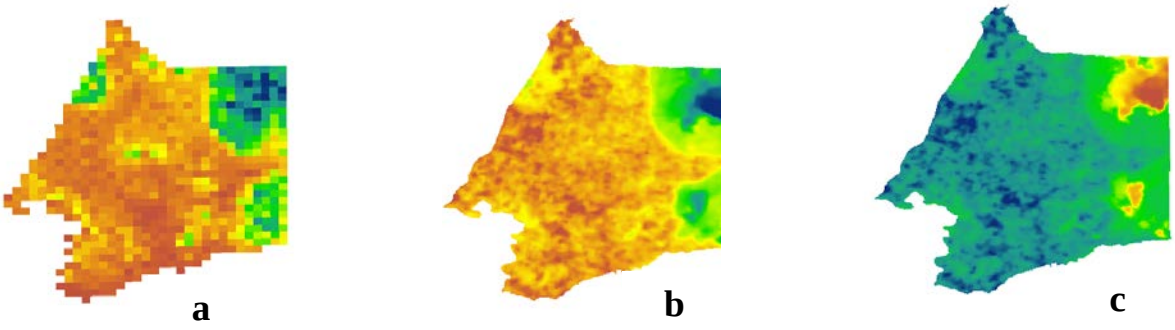


Figure 4.16a, b and c: Soil Organic Carbon, Land Surface Temperatures and Soil Moisture Index and results for Sorghum crop in Chiakariga study area

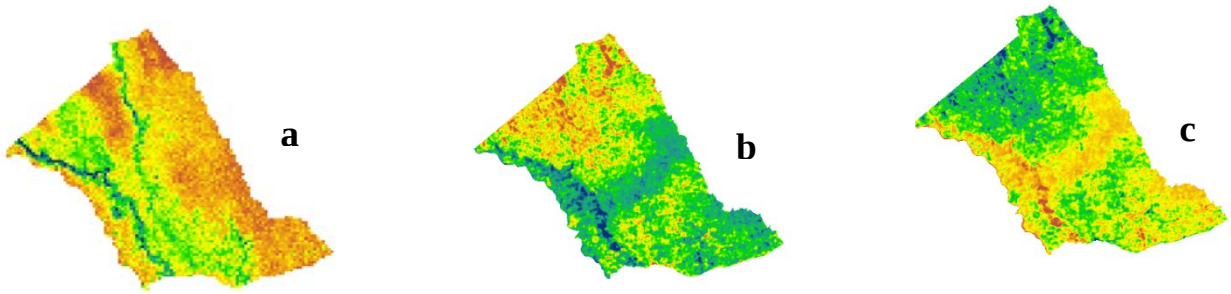


Figure 4.17a, b and c: Soil Organic Carbon, Land Surface Temperatures and Soil Moisture Index results for green gram crop in Ikombe-Katangi study area

4.3.3 Soil nitrogen, phosphorus and potassium results for green gram and sorghum study area

The extracted images for the soil nutrients for sorghum and green gram are shown in Figures 4.18 a, b, and c and 4.19 a, b, and c for the sorghum and green gram crops, respectively. The data extracted from satellite imagery estimation was correlated to the field data, and the results are shown in Tables 4.13–4.15 and Figures 4.20 a and b, 4.21 a and b, and 4.22 a and b. R^2 for the green gram crop was 0.92,0.99 and 0.86 for nitrogen, phosphorus, and potassium, while those for sorghum were 0.93,0.99 and 0.98 for nitrogen, phosphorus, and potassium, respectively. The highest and lowest values for estimated nitrogen, phosphorus, and potassium were recorded as 0.128–0.16 g/kg, 0.162–0.018 g/kg, and 135.2–98.2 g/kg, respectively, for the green gram crop, while for the sorghum study area, the estimated values for nitrogen, phosphorus, and potassium were recorded at 0.014–0.043 g/kg, 3.704–0.241 g/kg, and 119.77–44.37 g/kg, respectively.

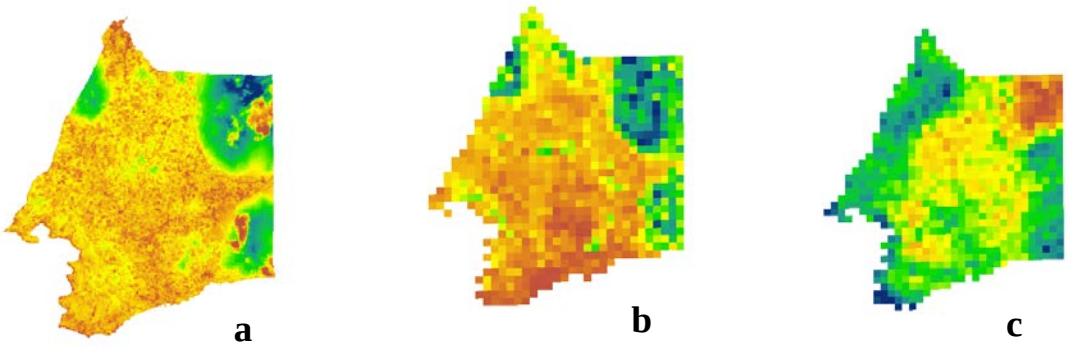


Figure 4.18a, b and c: Nitrogen, Phosphorus and Potassium results for Sorghum crop in Chiakariga study area

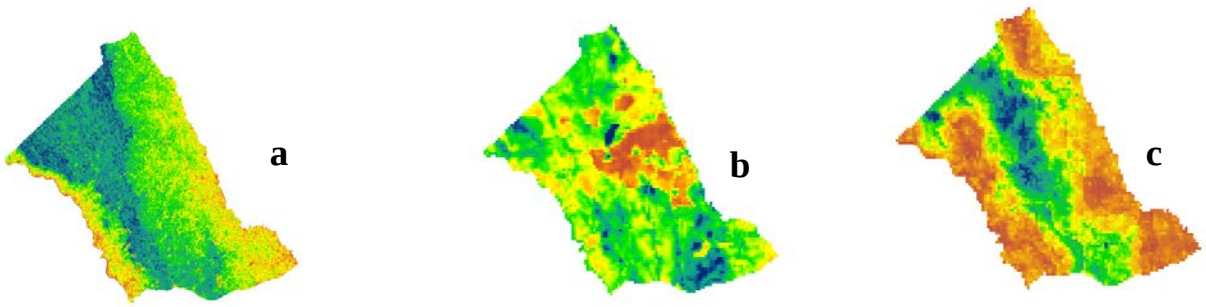


Figure 4.19a, b and c: Soil Nitrogen, Soil Phosphorus and Soil Potassium results for green gram crop in Ikombe-Katangi study area

Table 4.14: Field and remote sensing extracted Soil Nitrogen Data

Crop	ESTIMATED N (g/kg)	OBSERVED N (g/kg)
Green grams	0.128	0.121
Green grams	0.032	0.03
Green grams	0.032	0.029
Green grams	0.065	0.092
Green grams	0.086	0.096
Green grams	0.016	0.015
Green grams	0.036	0.028
Green grams	0.028	0.037
Green grams	0.036	0.035
Sorghum	0.013	0.006
Sorghum	0.043	0.052
Sorghum	0.014	0.02

Table 4.15: Field and remote sensing extracted Soil Phosphorus Data

CROP	ESTIMATED P (g/kg)	OBSERVED P (g/kg)
Green gram	0.036	1.308
Green gram	0.089	0.465
Green gram	0.096	0.778
Green gram	0.018	0.098
Green gram	0.036	4.637
Green gram	0.035	1.970
Green gram	0.042	3.488
Green gram	0.162	2.049
Sorghum	0.241	0.305
Sorghum	3.074	3.125
Sorghum	0.790	0.872

Table 4.16: Field and remote sensing extracted Soil Potassium Data

CROP	ESTIMATED K (g/kg)	OBSERVED K (g/kg)
	129	131.4
Green gram	121.7	124.8
Green gram	98.2	116
Green gram	107.4	118.3
Green gram	130.1	125.8
Green gram	100.1	114.2
Green gram	135.2	136.4
Green gram	100.9	106.6
Green gram	125.0	128
Sorghum	44.3712	47.5
Sorghum	93.5135	107.9
Sorghum	119.7665	123.8

Key nutrients nitrogen (N), phosphorus (P), and potassium (K) play a major role in global food production (Smil, 2000; Steen, 1998). Soil fertility in a study by Kihara *et al* (2020). highlights three key nutrients: nitrogen (N), phosphorus (P), and potassium (K), which play a major role in global food production. When addressing the need for plant nutrients with regard to critical global food problems, it is essential to know the quantities of nutrients required for target crop production and the supply of nutrients by the soils used for crop growth. In this case, the study estimated the levels of N, P, and K at 0–40 cm and validated them with field data in Tables 4.14–4.16. The R2 was 0.9276, 0.9949, and 0.8558 for N, P, and K in the green gram crop study areas. For sorghum, the R2 was 0.9282, 0.9986, and 0.9773 for N, P, and K in

the sorghum study area farms. The validations indicate a high correlation between the field data and estimated satellite imagery data.

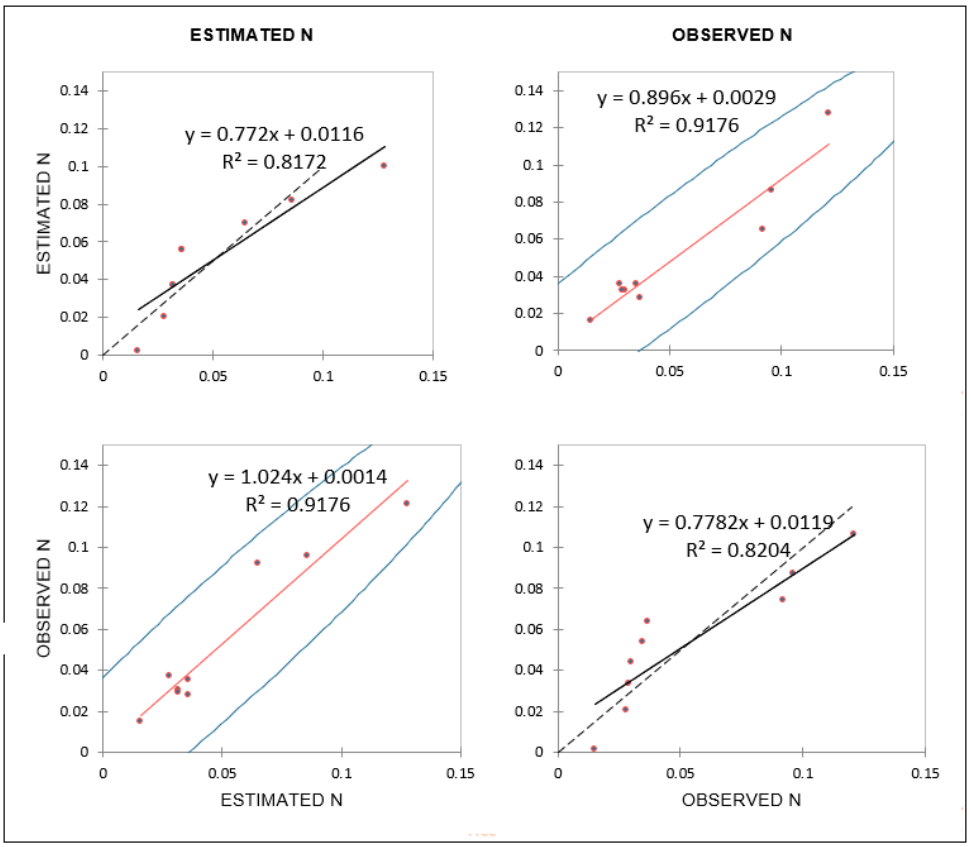


Figure 4.20a: Regression results for remote sensing-based Soil Nitrogen correlations with green gram field data

Soil nitrogen has been found to be important in defining soil quality and improving soil fertility, which affects crop production (Iodice *et al*., 2021). The high correlation with lab soil sample analysis, as shown in Figure 4.20a and b, which is an indication that spatial analysis of soil nitrogen is possible. A study by Cole *et al*. (2023) used remote sensing data to extract soil nitrogen data for monitoring and improving fertilizer use among smallholder farmers in India. Mashaba-Munghemezulu *et al*. (2021) modeled the spatial distribution of soil nitrogen using machine learning technologies for improved soil nutrient management among smallholder farmers. The study concluded that developing frameworks for the mapping of spatial variability of

soil nitrogen is important for local governments, farmers, and other stakeholders in identifying areas of excess or deficiencies.

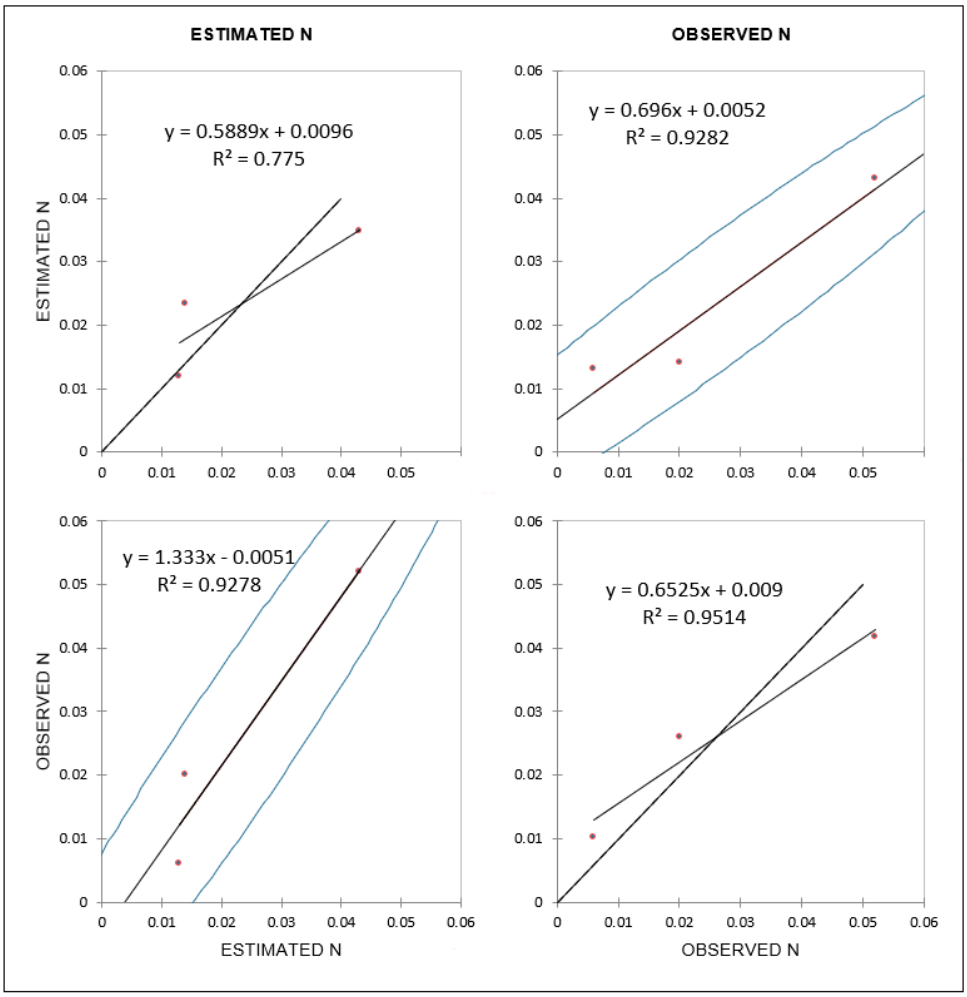


Figure 4.20b: Regression results for remote sensing-based Soil Nitrogen correlations with sorghum field data

Soil nitrogen in sorghum productionhas been found to be very important for good yields. Nutrient management is one of the most productive ways to increase sorghum yields. Nitrogen tends to be the most limiting nutrient in crop production (Fageria and Baligar, 2005; Gweyi, 2006; Gweyi-Onyango *et al*., 2009). It takes the correct timing and amount of nitrogen to maximize sorghum grain yield. A lack of these vital nutrients at the critical growth stages will negatively affect sorghum yield. The

importance of soil nitrogen for sorghum yield production was carried out by Davidson (2019) and found to be successful.

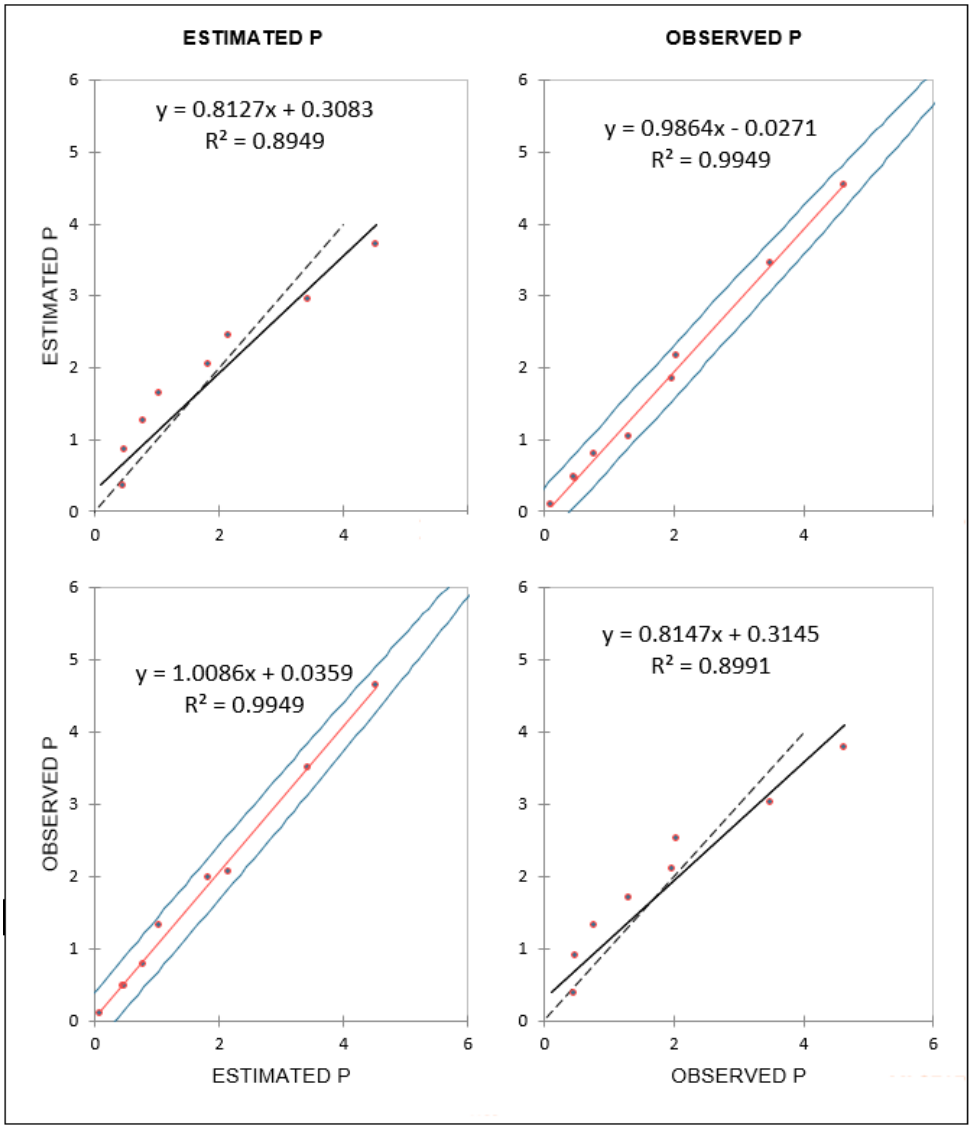


Figure 4.21a: Regression results for remote sensing-based Soil Phosphorus correlations with green gram field data

There was a high correlation between the satellite-based soil phosphorus and soil sample analysis for soil phosphorus, as shown in Figure 4.21a and b for green gram and sorghum, respectively. An indication in use of the spatial distribution of soil phosphorus for monitoring of cop development and data estimation. A large amount

of P is required for metabolic pathways of energy transfer that take place during nodule functioning (Mitran *et al.*, 2018). But most of the agricultural soils have inadequate amounts of P to support efficient nodule development in legumes such as green grams. On the other hand, soil P has been found to affect all parts of the sorghum crop, from the development of the rooting system to maturity and grain filling (Spence and Welch, 1999).

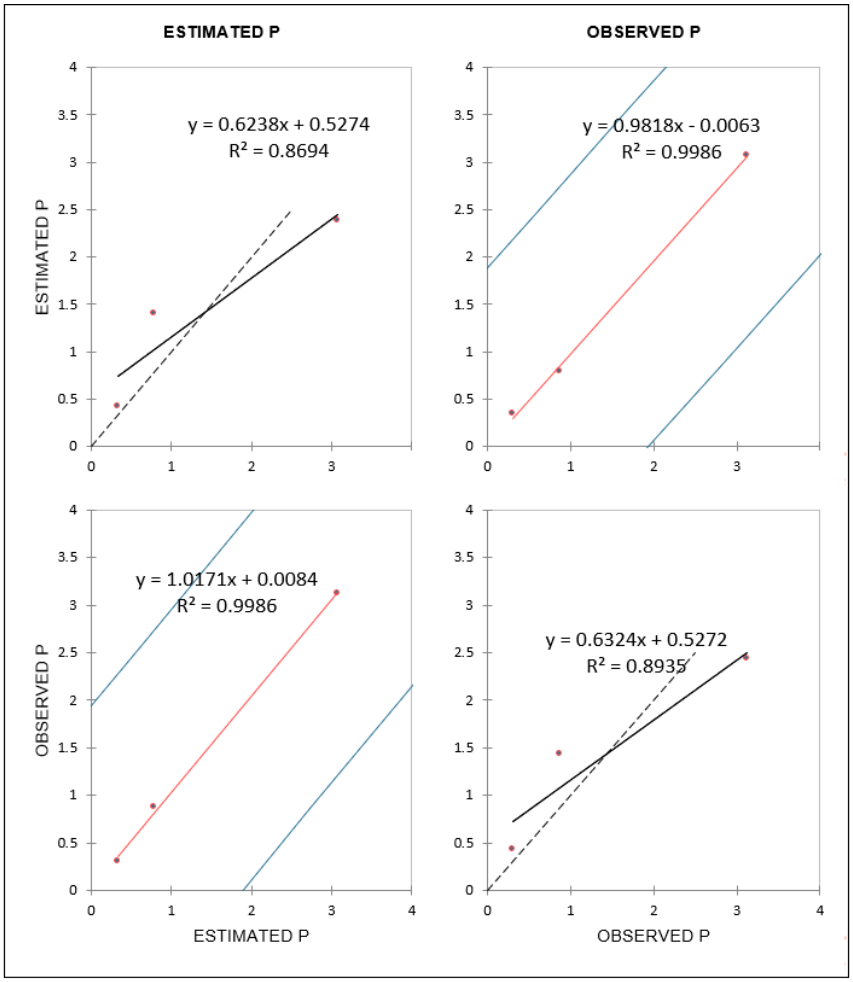


Figure 4.21b: Regression results for remote sensing-based Soil Phosphorus correlations with Sorghum field data

Spatial analysis and distribution of soil p were successfully done by Kim *et al*. (2014). The study was able to confirm that prediction of soil P was possible in wetland soils at different spatial resolutions. Another study by Ray *et al*. (2002) showed the

effectiveness of high-resolution data in estimating soil nutrients-related parameters and the development of field variability useful in crop management. The study, however, pointed out the challenge of the gap between satellite data acquisition and field measurements. The two studies confirm the possibilities of generating soil P from remote sensing data with high precision.

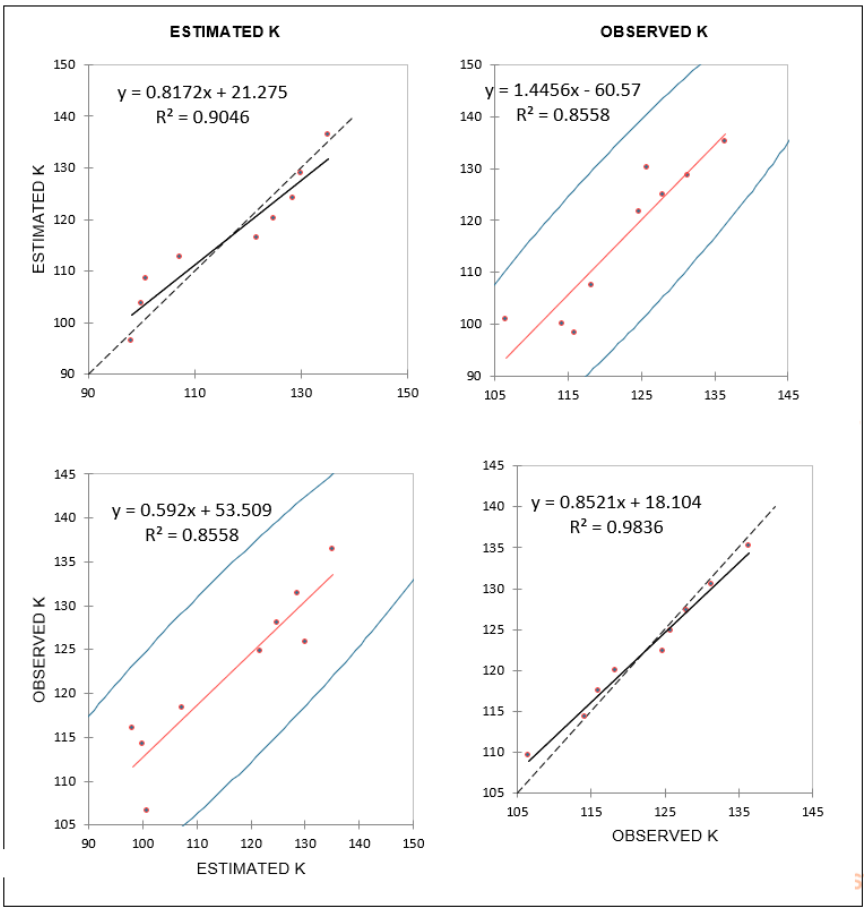


Figure 4.22a: Regression results for remote sensing-based Soil Potassium correlations with green gram field data

The availability of potassium (K) has been found to affect photosynthesis, water use efficiency, plant tolerance to diseases, and drought (Singh, 2017). Pratap Singh (2017), attributed growth and yield characters to being affected by the levels of K in green gram. The correlation between satellite-extracted K and soil sample-measured K, as shown in Figure 4.22 a and b for both green gram and sorghum, was very high.

Soil K has been found to affect root growth and nutrient uptake in both sorghum and green gram (Fageria, 2009). (Jat *et al*., 2022) indicate that the significant application of soil K resulted in an increase in yield and protein content in green gram. This confirms the importance of soil K in crop data estimation in this research study.

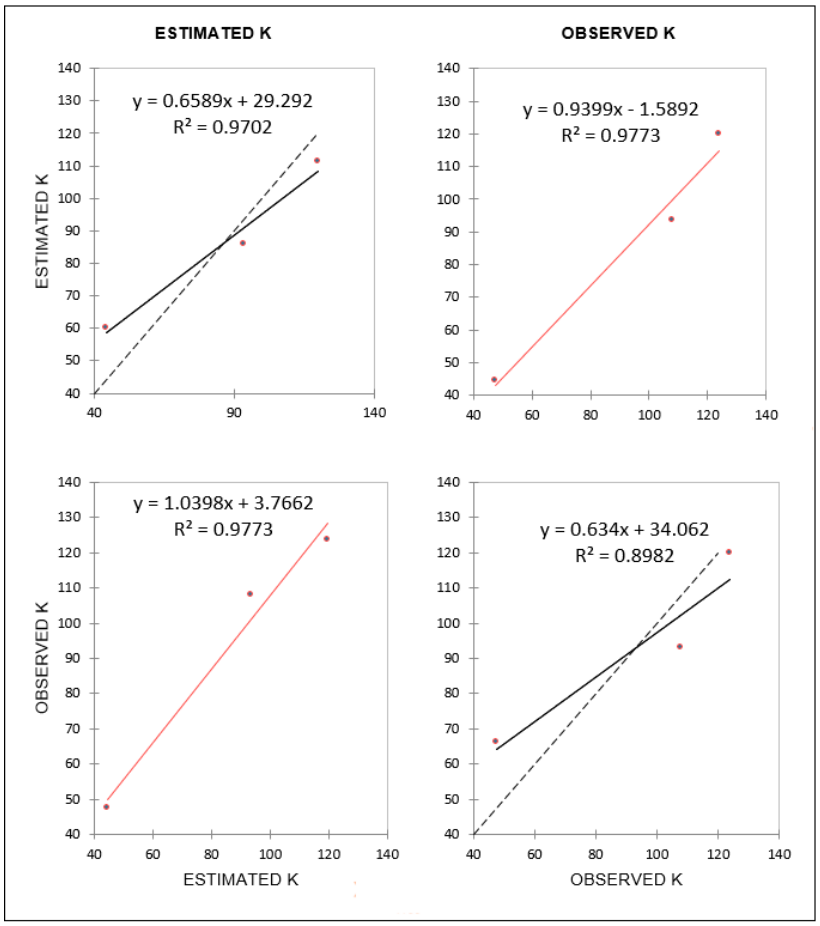


Figure 4.22b: Regression results for remote sensing-based Soil Potassium correlations with sorghum field data

Spatial variability of soil K was evidently found by Mwendwa *et al*. (2022) in their research on soil nutrient spatial variability in coffee farms in Kabete, Kenya. Landsat 8 soil-based prediction models for soil K performed well in identifying the nutrient status in the soil (Xu *et al*., 2017). The data being free could find a wide range of use in field-specific soil management aspects for smallholder farmers (Xu *et al*., 2017). This confirms the possibility of extracting soil K using satellite data.

4.3.4 Evapotranspiration, soil pH and rainfall results for green gram and sorghum study areas

The extracted evapotranspiration, soil pH, and rainfall are presented in Figure 4.23 a, b, and c for sorghum, while for green gram, they are presented in Figure 4.24 a, b, and c, respectively. The result for the estimated evapotranspiration for green gram and sorghum was 2.36-3.82mm/day and 2.17-2.73mm/day, respectively, while rainfall was an average of 3.81mm/day for green gram and sorghum. Average evapotranspiration is estimated at 3.05 mm/day and 2.47 mm/day for green gram and sorghum, while average rainfall is recorded at 3.81 mm/day and 9.9 mm/day. According to the FAO, the Penman-Monteith (PMon) method gives more consistent evapotranspiration estimates and has been shown to perform better than other methods (Monteith, 1965; Monteith, 1981; Palaskar *et al*., 1987). Spectral analysis through the extraction of important data such as evapotranspiration was found to enable the analysis of crop growth, photosynthetic size, water use, and crop yield as affected by environmental factors and stresses (Wiegand and Richardson, 1990). Evapotranspiration data is key in determining water use efficiency in crops, especially in semi-arid areas. Evapotranspiration was found to be very critical in the determination of crop data such as yield in green gram and sorghum that are commonly grown in semi-arid areas. Historically, crop stress has been assessed through the utilization of vegetation-based indices, precipitation, and soil moisture (Marshall *et al*., 2012). The ET is more closely associated with moisture availability to stressed crops than precipitation or volumetric soil moisture; however, due to inadequate parameterization and a general dearth of calibration and validation data, it has not been extensively utilized in the past (Marshall *et al*., 2012).

The rainfall estimate agrees with what Mugo *et al*. (2016) used in the research. Rainfall amounts are critical for crop growth and determine the crop yield in sorghum and green gram.

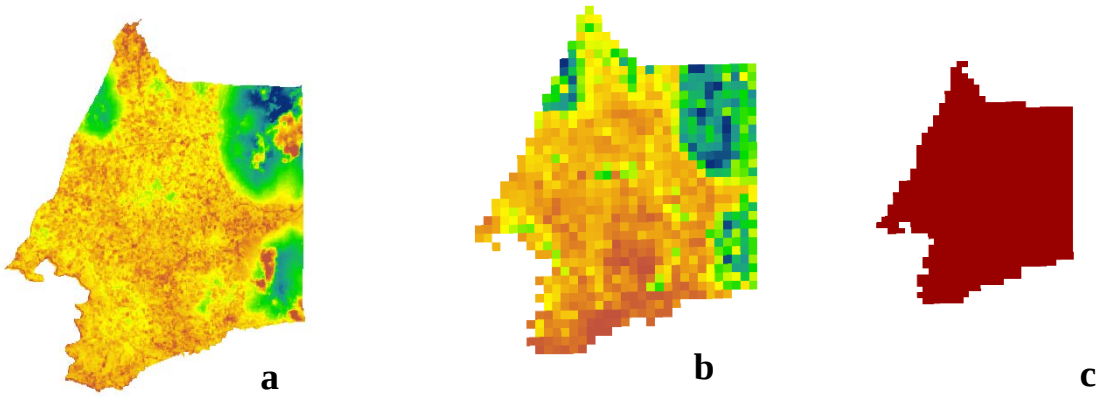


Figure 4.23a, b and c: Evapotranspiration, Soil pH and Rainfall results for Sorghum crop in Chiakariga study area

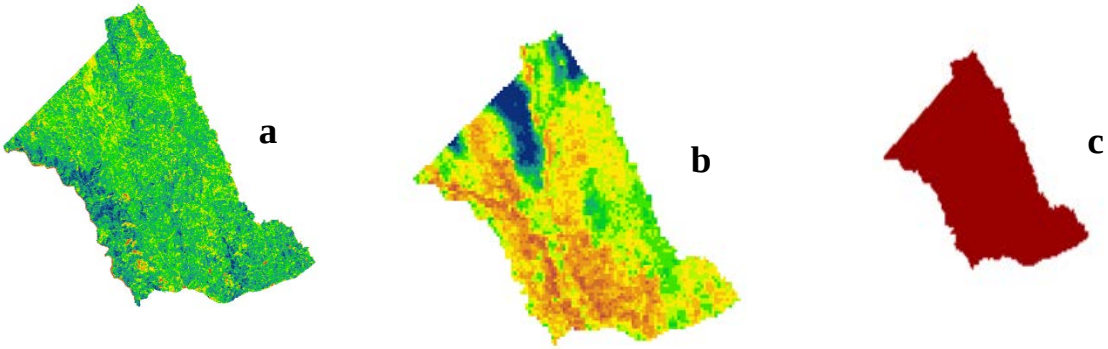


Figure 4.24a, b and c: Evapotranspiration, Soil pH and Rainfall results for green gram crop in Ikombe-Katangi study area

In their research, Tyagi *et al.* (2000) found sorghum to have a minimum evapotranspiration of 3 mm/day, peaking at 6 mm/day during the reproductive stage

and going down to 4 mm/day during the reproductive stage. These studies confirm the results obtained from the field on the evapotranspiration rates for green gram and sorghum in the Ikombe-Katangi and Chiakariga study areas. It is important to point out that modern drought monitoring systems combine evapotranspiration, soil moisture, and even vegetation anomalies to track drought status.

Table 4.17: Field and remote sensing extracted Soil pH Data

Crop	Observed soil PH	Estimated soil PH
Green grams	6.5	6.4
Green grams	7.1	6.4
Green grams	7.8	6.8
Green grams	7.2	7.5
Green grams	7.5	7.0
Green grams	7.0	7.0
Green grams	7.9	6.9
Green grams	7.8	7.6
Green grams	7.8	7.8
Sorghum	7.8	6.8
Sorghum	7.6	6.5
Sorghum	7.9	7.0

The soil pH estimated data was validated using field data, and the results are presented in Figure 4.50a and b for the green gram and the sorghum crop, respectively. Such data are in concurrence with those of Manzi *et al*. (2023a) in the same region. The data used in the validation is shown in Table 4.17. The R2 for green gram for the soil pH was 0.9141, while that for sorghum was 0.9944, showing a high correlation with the field data. The soil pH range for the green gram study area was estimated at 6.4–7.8, while that of the sorghum study area was 6.5–7.5. According to the study by Imalasiri *et al*. (2020), Soil pH was found to be important in the MONICAH agro-ecosystem model for yield estimation. It was noted that soil pH and other parameters

affect crop growth. The values of soil pH were found to have a high correlation with the field data, as shown in Figure 4.25a and b. Rainfall has emerged as very critical in the determination of crop yield (Makowski *et al*., 2006). The average extracted rainfall amount data for the study area was 3.81 mm/day and 9 mm/day for the green gram and sorghum study areas, as shown in Figures 4.23 c and 4.24 c.

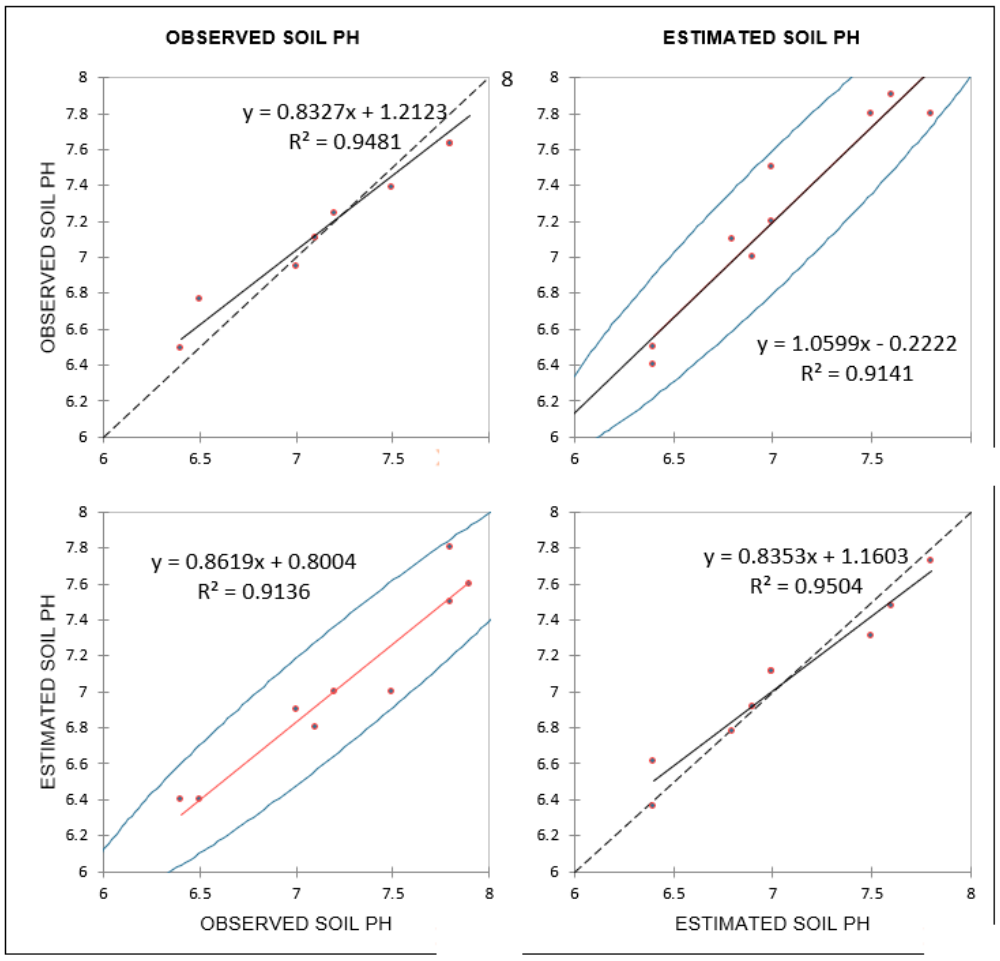


Figure 4.25a: Regression results for remote sensing-based Soil pH correlations with Green gram field data

Soil pH affects the accessibility of specific nutrients and micronutrients, phosphorus in particular, in addition to biological activity. The soil pH and salt composition are influenced by various factors, including the age of the soil, the initial soil parent material, the category of plants cultivated, and climatic conditions, particularly

rainfall volume (Natarajan *et al*., 2022). Soil pH is a significant quality indicator variable due to the fact that it regulates numerous chemical and biological processes occurring in soil. The measurement of soil pH, which indicates acidity or alkalinity, is of utmost importance in crop management as it regulates the availability of nutrients to the crop (Natarajan *et al*., 2022).

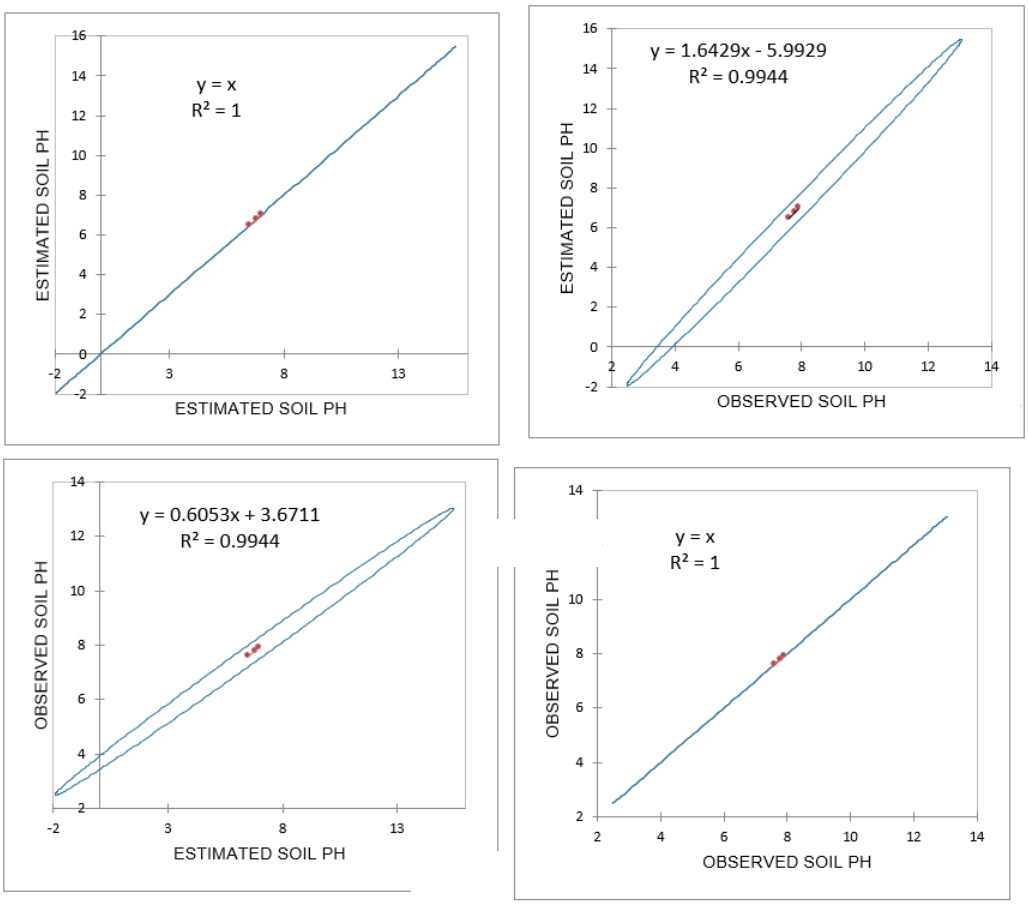


Figure 4.25b: Regression results for remote sensing-based Soil pH correlations with sorghum field data

Ghazali *et al*. (2020) observed hat satellite data with high spatial, spectral, and temporal resolutions can estimate soil pH with fairly good accuracy. This confirms the use of Landsat satellite data in the extraction of soil pH for crop data estimation in green gram and sorghum, as shown in Figure 4.24b. This is also confirmed by the

high correlation between the satellite data and soil sample data, as shown in Figure 4.25a and b for green gram and sorghum, respectively.

4.3.5 Sorghum and green grams yield model under farm field conditions

The twelve parameters analyzed were run through a random forest machine learning algorithm to generate the yield estimates for sorghum. The validation of the model was done using root mean square error (RMSE) and root mean absolute error (RMAE), and the results are presented in Tables 4.18 to 4.19 for green gram and sorghum, respectively. The field estimates generated are correlated to field data, and the results are shown in Tables 4.18 and 4.19 for green gram and sorghum. The RMSE was 4.036 and the R2 was 0.98, while the RMAE was 3.022, for the green gram crop, and the RMSE was 6.51 and the R2 was 0.99, while the RMAE was 5.5.

Table 4.18; Farm field yield estimation of green gram crop at the vegetative stage during the October November December rains season for Ikombe-Katangi study area

FARM	OBSERVED YIELD(Kg/Acre)	ESTIMATED YIELD(Kg/Acre)	ESTIMATED - OBSERVED
Green gram	40	38	-2
Green gram	90	100	10
Green gram	5	1.5	-3.5
Green gram	70	69	-1
Green gram	10	12	2
Green gram	8	6.8	-1.2
Green gram	40	38	-2
Green gram	120	115.7	-4.3
Green gram	25	23.8	-1.2
Average	**45.3**	44.9	
r	**0.994828556**		
R2	**0.989683855**		
RMSE	**4.036224859**		
RMAE	**3.022222222**		

Table 4.19; Farm field yield estimation of sorghum crop at the vegetative stage during the October November December rains season for Chiakariga study area

FARM	OBSERVED YIELD (Kg/Acre)	ESTIMATED YIELD (Kg/Acre)	ESTIMATED-OBSERVED
Sorghum	100	98.5	-1.5
Sorghum	450	445	-5
Sorghum	1950	1940	-10
Average	**833.3**	827	
r	**0.999999448**		
R2	**0.999998896**		
RMSE	**6.512807894**		
RMAE	**5.5**		

The scatter plot for correlations between estimated and observed datasets for green gram and sorghum is presented in Figure 4.26 a and b. From the results, it can be confirmed that low levels of yield have been attributed to poor rainfall, soil moisture stress, and low nutrient availability. These have had a major constraint that impinges crop productivity in arid and semiarid environments around the world (Yazar and Ali, 2016). Although remote estimation for crop yield has been studied for decades and various models have been developed and worked well in selected experimental areas, no routine or universal model is applicable for a wide range of operational applications and for all crop species (Ferencz *et al*., 2004; Yu *et al*., 2019).

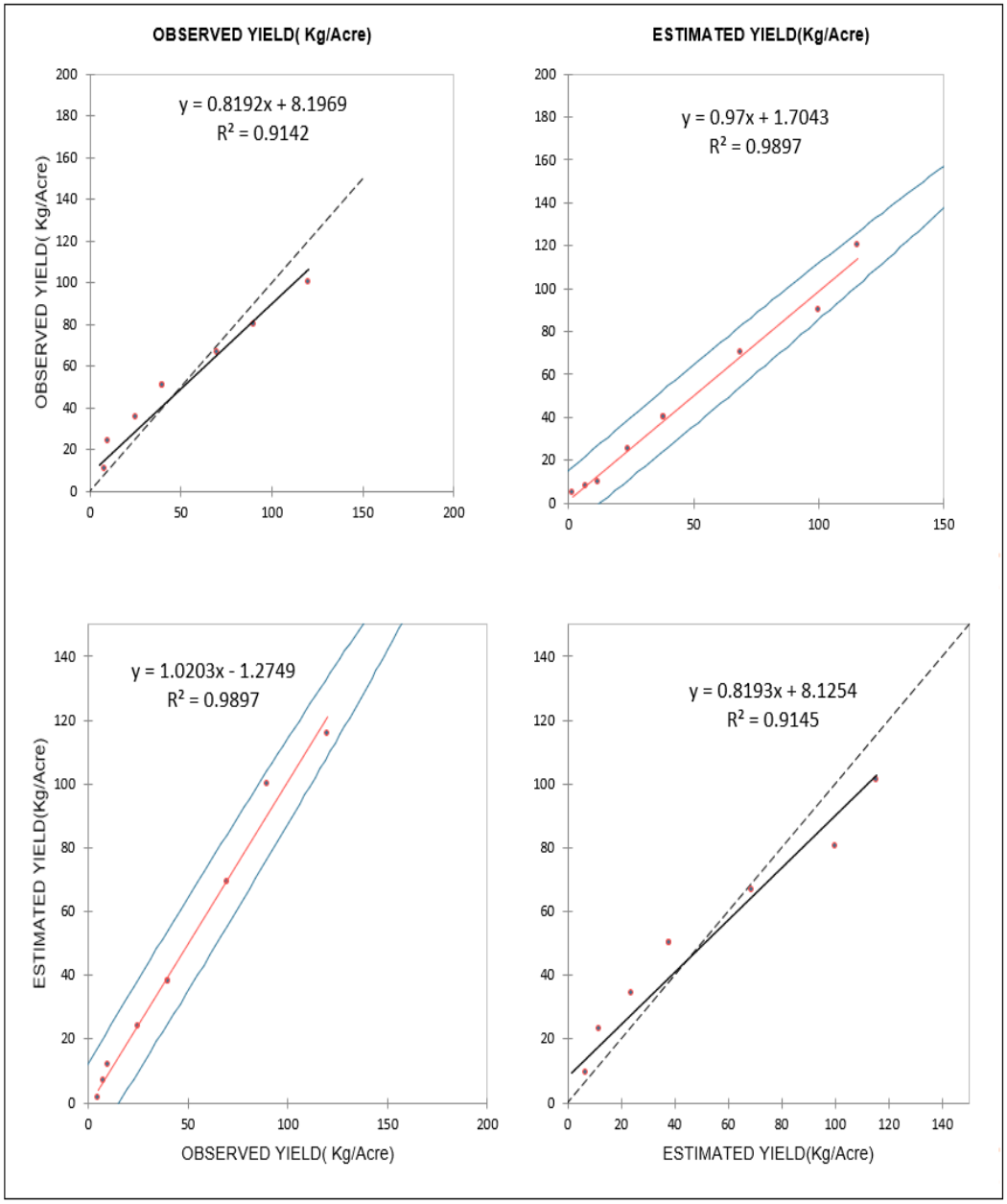


Figure 4.26a: Regression correlation of the estimated satellite imagery data to green gram field data.

A high correlation between the estimated crop yield from satellite data and the observed field yield was evident in the research study for the green gram and sorghum crops, as illustrated in Figure 4.26a and b. This confirms the ability to use earth observation for cop data estimation under farm field conditions.

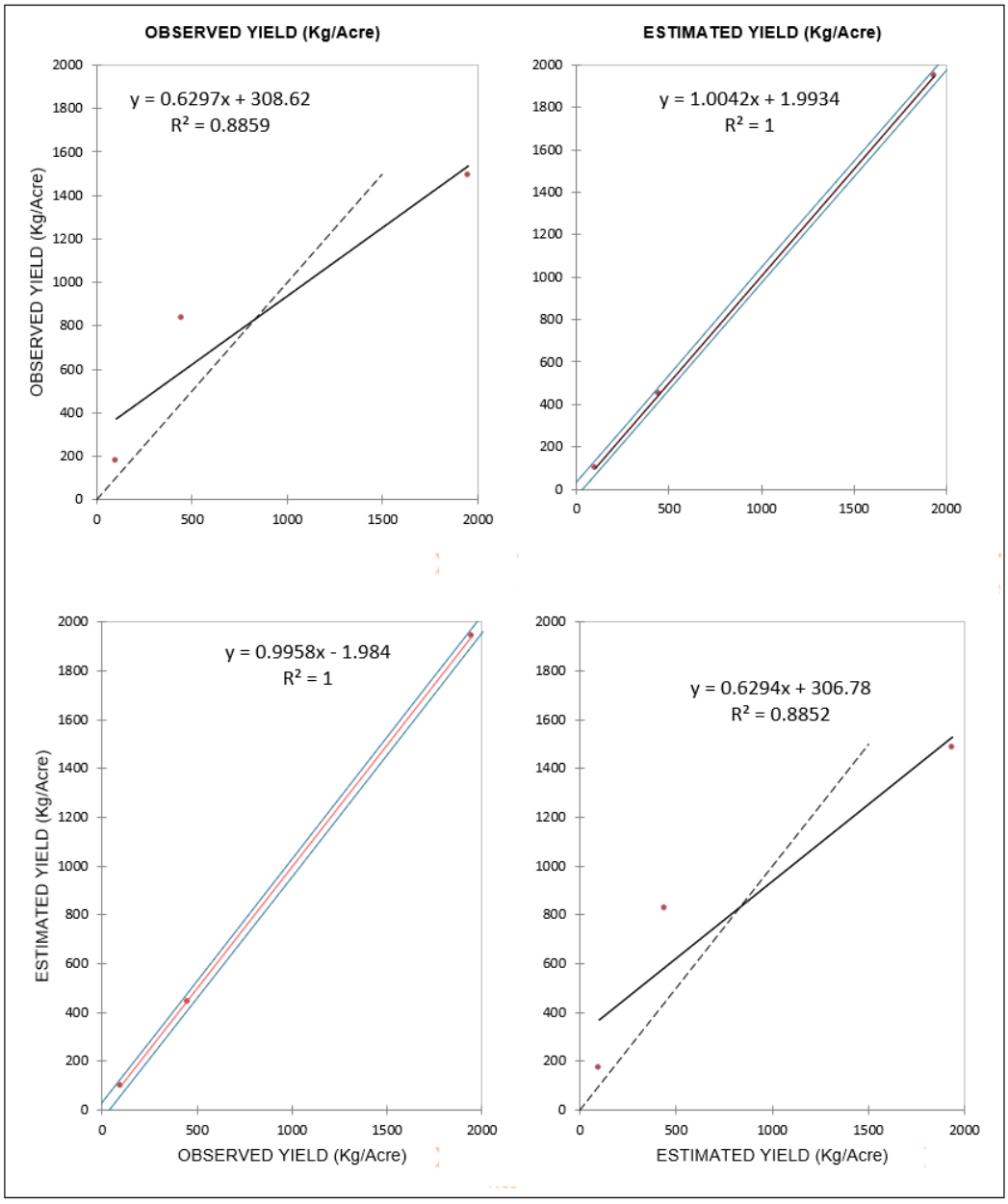


Figure 4.26b: Regression correlation of the estimated satellite imagery data to sorghum field data.

Many models can achieve high accuracy of yield prediction at quite large scales, such as county- or country-levels, but may not refine the detailed variations at the field level, especially for small-size croplands planted with different species or under various field managements (Lobell and Gourdji, 2012). This situation may commonly

exist in many developing countries where smallholder croplands mainly occupy cultivated lands (Roy *et al*., 2016). The current study developed a dynamic model based on ten parameters suitable to assess farm field condition yield estimation of green gram and sorghum crops. Through weighted least squares estimation, the weight of each parameter was determined. Traditionally, crop yield estimation is dependent upon data collection techniques from ground-based field visits. Such techniques are often subjective, costly, and prone to large errors, leading to poor crop assessment and crop area estimation (Reynolds *et al*., 2000).

Remote sensing data has the potential to provide timely, systematically high-quality, spatially accurate information about land features, including environmental impacts on crop growth (Liu and Huete, 1995). The temporal dynamics of remote sensing data and their close relation to plant characteristics could play a crucial role in establishing an effective pre-harvest yield estimation method. The most multispectral satellite systems measure varies spectral bands within the visible to mid-infrared region of the electromagnetic spectrum (Shwetank *et al*., 2010). The development of green gram and sorghum dynamic models for field-condition crop yield estimation in small-holder farms adds to the possibilities of remote sensing-based yield estimation models. The results of crop yield were compared to field results, and evidently, the R^2, RMSE, and RMAE errors were 0.99, 4.036, and 5.5 for green gram, indicating a high probability of the model being able to estimate the crop yield, as shown in Table 4.18. On the other hand, the R^2, RMSE, and RMAE for sorghum were 6.51, 0.99, and 5.5, respectively, indicating moderate probabilities in the model's capabilities in estimating the crop yield, as shown in Table 4.19. The Sorghum crop model was affected by limited field data for the calibration and validation of the model. The most common practice to evaluate a model is to compare observed data versus estimated data (outputs) using a metric or indicator that measures the distance between these observed and simulated data (Wallach *et al*., 2019). Various metrics exist for models in general, and many of these have been transferred for use in evaluating spatialized crop models.

4.4 GIS based data estimation tool for Sorghum and green grams under farm field conditions

The dynamic model development was customized through ARCGIS model builder software to enable ease of usability by different users. The process developed four models, which is a step-wise process in crop data estimation. The first model tool developed was for farm area estimation and digitization, which allows one to operate within the farm boundary while farm data is estimated while at the same time enabling automation of spectral signature identification through the pre-loaded spectral signature libraries. This process enhances crop identification and increases the accuracy levels of crop data estimation (Figures 4.27 and 4.28). The second process led to the development of an agroecological sub-zonation model tool that ensures a homogenous environment for farm field data estimation (Figures 4.29 and 4.30). The last model tool is the yield crop estimation, which preloads all required remote sensing-prepared data for the yield estimation for green gram crop and sorghum. The developed model and tools are shown in Figures 4.31–4.32. A combination of these models was packaged and put into Arctool in ARCGIS software, as shown in Figure 4.33. To use traditional point crop models on bigger areas, spatialization of crop models for data estimation is used. This has been done at the field scale (Acevedo-Opazo *et al*. 2010), the multiple field scale (Baralon *et al*. 2012), the regional scale (Balkovič *et al*. 2013), the continental scale (Adam *et al*. 2011), and the regional scale (Battude *et al*. 2016). This study spatialized a crop model for use at the agroecological zone level while ensuring a homogenous environment. Further, stepwise model execution enables the use of multi-model ensembles that reduce the estimation error. Using GIS and remote sensing data, spatialized crop models can tell the difference between crop classes, guess the yield (Panda *et al*., 2010), and take into account differences in space, which makes field scouting more efficient (Schuler, 2002). Using multi-model ensembles (MMEs) is a new approach in crop modeling for data estimation (Wallach *et al*., 2019). These types of models tend to decrease the estimation error (Wallach *et al*., 2018). These MMEs allow for increased accuracy in crop data estimation (Martre *et al*., 2015).

This study has combined several models in addition to the crop model to improve on the accuracy in the estimation of crop data.

4.4.1 Farm delineation and spectral signature range model

The farm delineation and spectral signature tool is developed based on pre-processing signature libraries, which are preloaded into the tool to aid in crop identification once the satellite imagery is loaded as part of the data input. Panda *et al*. (2010) demonstrated the importance of GIS in the development of models that can enable crop identification through spectral reflectance. Their study discusses the crop-specific crop management system that had an inbuilt model for the extraction of spectral reflectance. Similarly, Panda (2003) demonstrated the importance of remote sensing data mining through a model for crop identification. The study utilized various ARCGIS models for use in extracting spectral reflectance data for identification of sorghum and green gram as part of the multiple models for data estimation for the study area. Figures 4.31 and 4-32 illustrate the front-end and back-end parts of the tool. In addition, the farm boundaries are also digitized and preloaded as part of the farm area estimation and identification procedure.

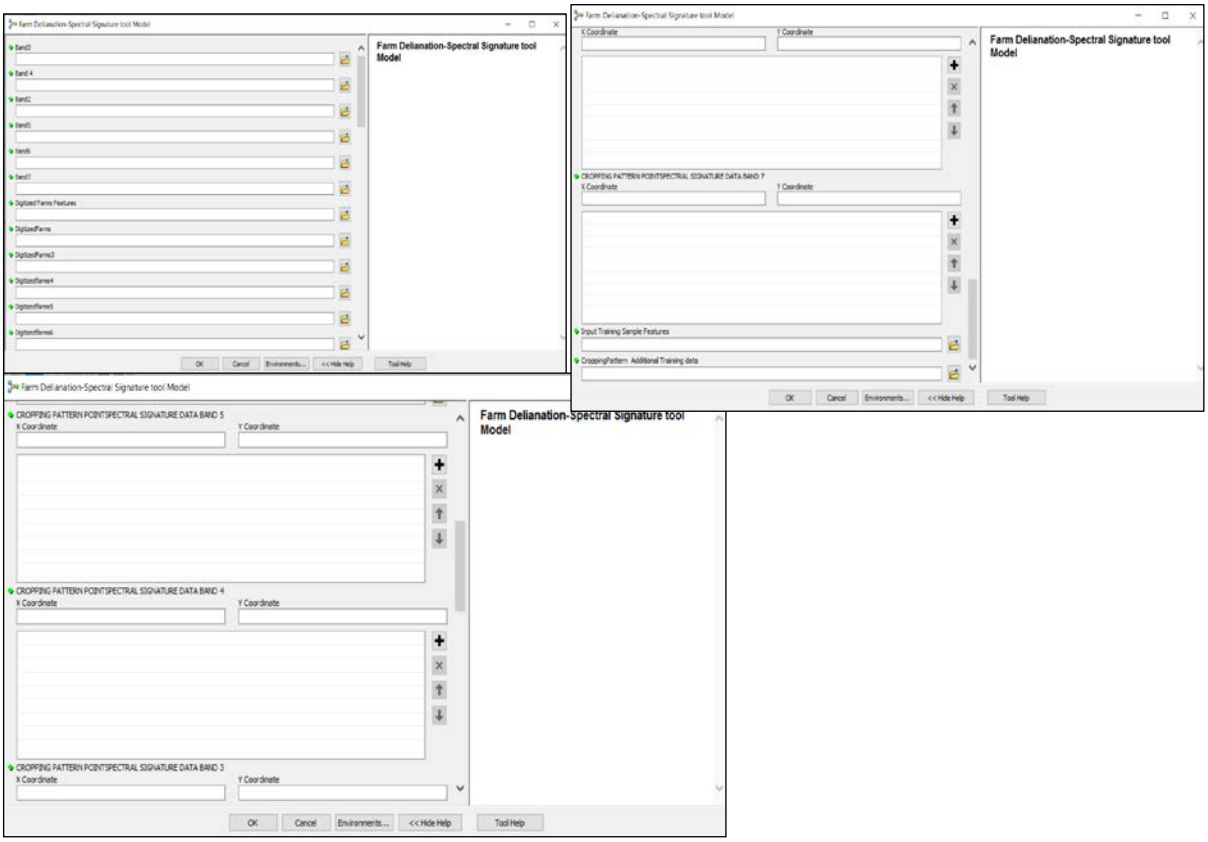


Figure 4.27: Front end part of the Farm delineation and spectral signature range model

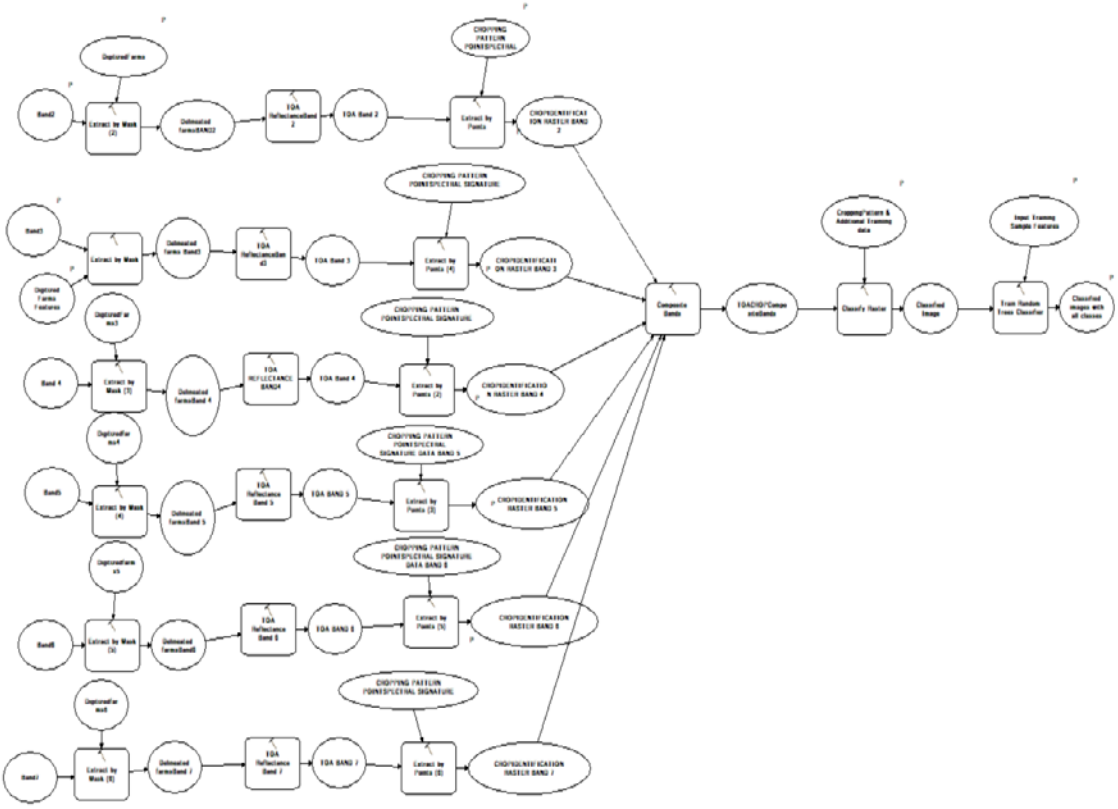


Figure 4.28: Back-end part of the arm delineation and spectral signature range model

4.4.2 Agro-ecological Sub-Zonation model

This model tool is designed to allow for the necessary data inputs for sub-zonation of agroecological zones as prescribed under the Materials and Methods chapter. Assessment of agri-environmental indicators using remotely sensed images was found to be very critical in crop identification as well as yield estimation (Torres *et al*., 2008). The importance of rainfall patterns in understanding and creating a homogenous environment for yield estimation has been emphasized by Panda (2008). Similarly, this study adopted several parameters and developed a model that analyses micro-climatic zones with the aim of creating homogenous areas. Once all the necessary datasets are put into the model, The final output is the identification of all sub-zones within the AEZ to allow for a homogenous environment for yield estimation at farm field conditions. Caldiz *et al*. (2001) modelled using GIS potential

agro-ecological zones for potato production in Argentina. The back end and front of the tool are illustrated in Figures 4.29 and 4.30 below.

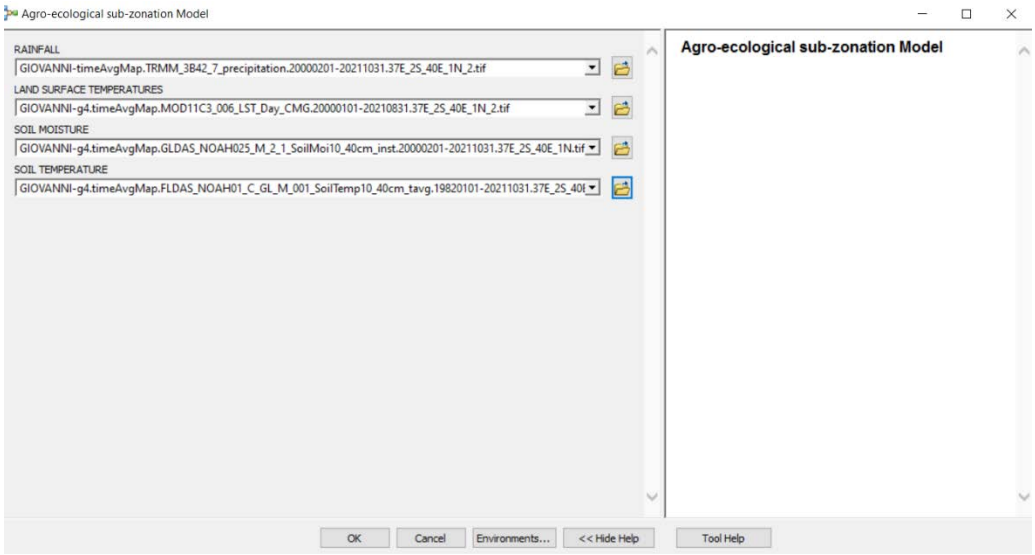


Figure 4.29: Front end part of the a Agro-ecological Sub-Zonation model

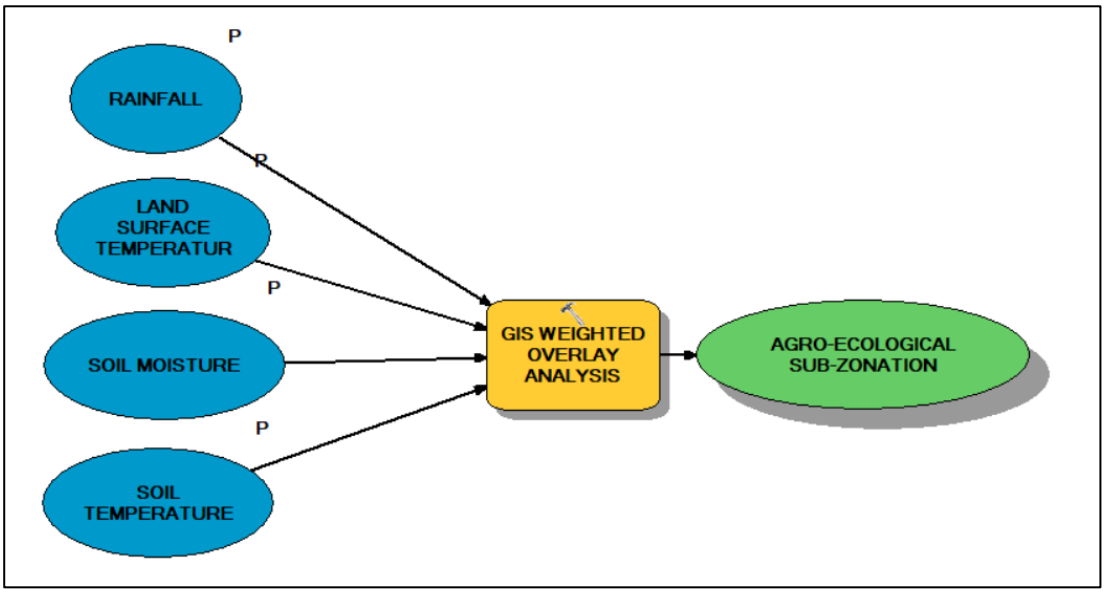


Figure 4.30: Back-end part of the agro-ecological Sub-Zonation model

4.4.3 Crop yield estimation model

The model tool is a combination of all the data input or the ten parameters involved in yield estimation undertaken in the current study. Liu *et al*. (2011) developed a yield estimation model that showed soil nutrition and precipitation to be the most important factors in oat production. The current study builds a model using ARCGIS to incorporate ten parameters that were found to be very key in green gram and sorghum production. The tool is designed to allow raw satellite imagery data entry since parameter preprocessing and extraction are built into the model. In a similar study, Wan *et al*. (2008) developed a GIS-based support system for monitoring oats in China. Their study confirmed the use of remote sensing data for inbuilt models for crop monitoring. Apart from satellite data, other prerequisite datasets, such as farm boundaries, would be required as an addition for the model to run (Panda 2010). An understanding of the various coefficients involved in parameter extraction is also necessary (Caldiz *et al*., 2001). Figures 4.31 and 4.32 illustrate the back-end process of the yield estimation model tool.

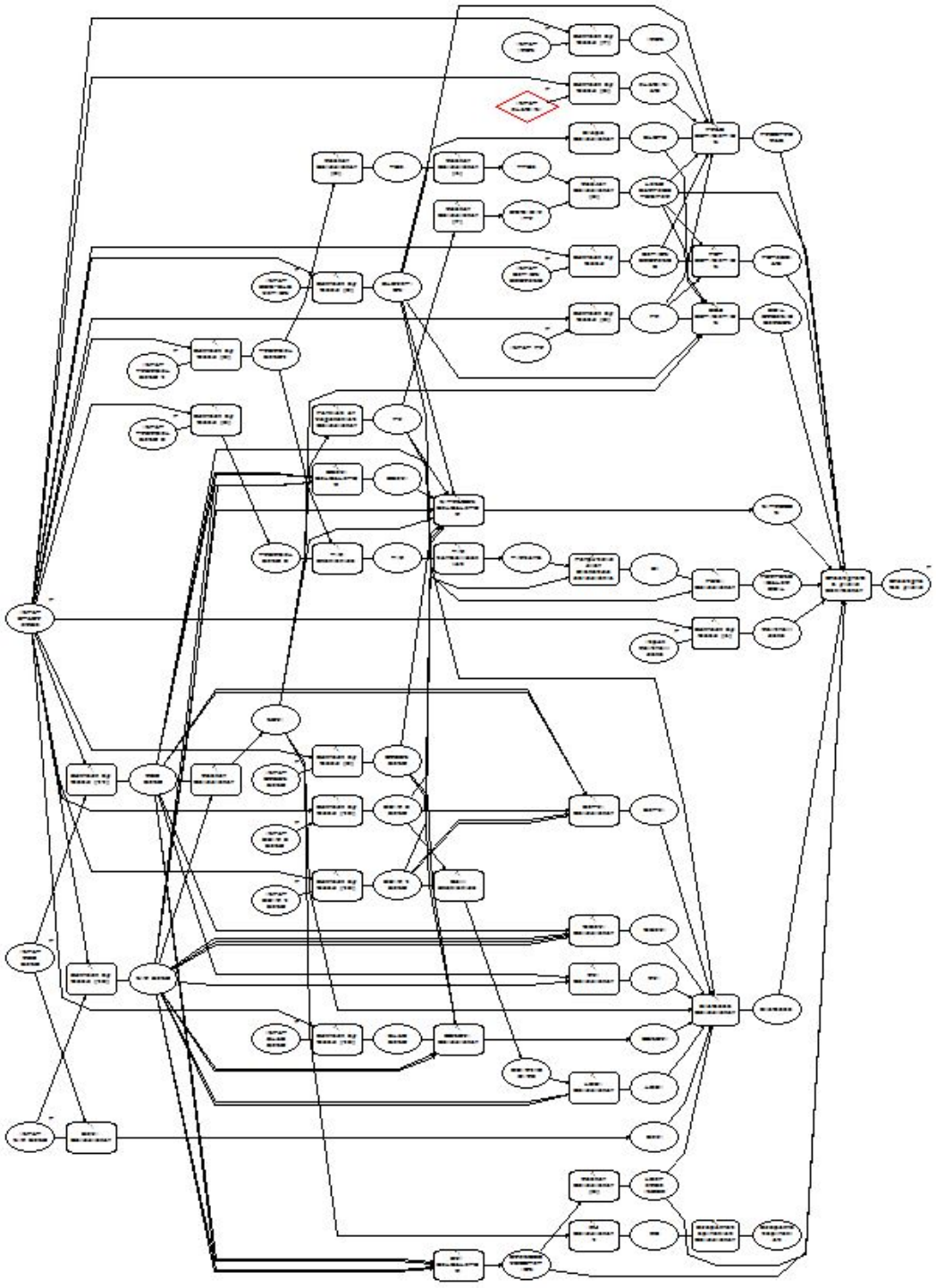


Figure 4.31: Back-end part of the yield estimation model for green gram

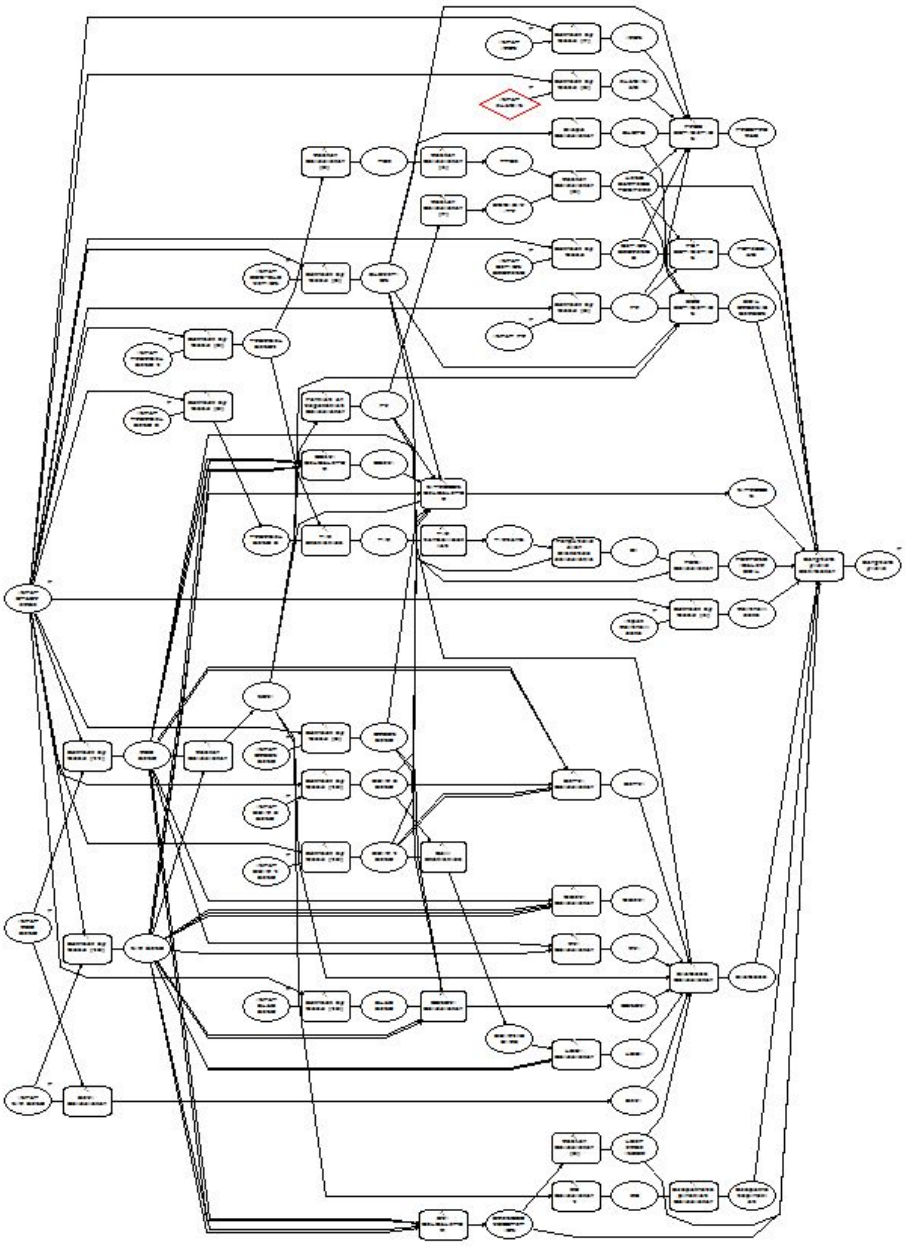


Figure 4.32: Back-end part of the yield estimation model for Sorghum

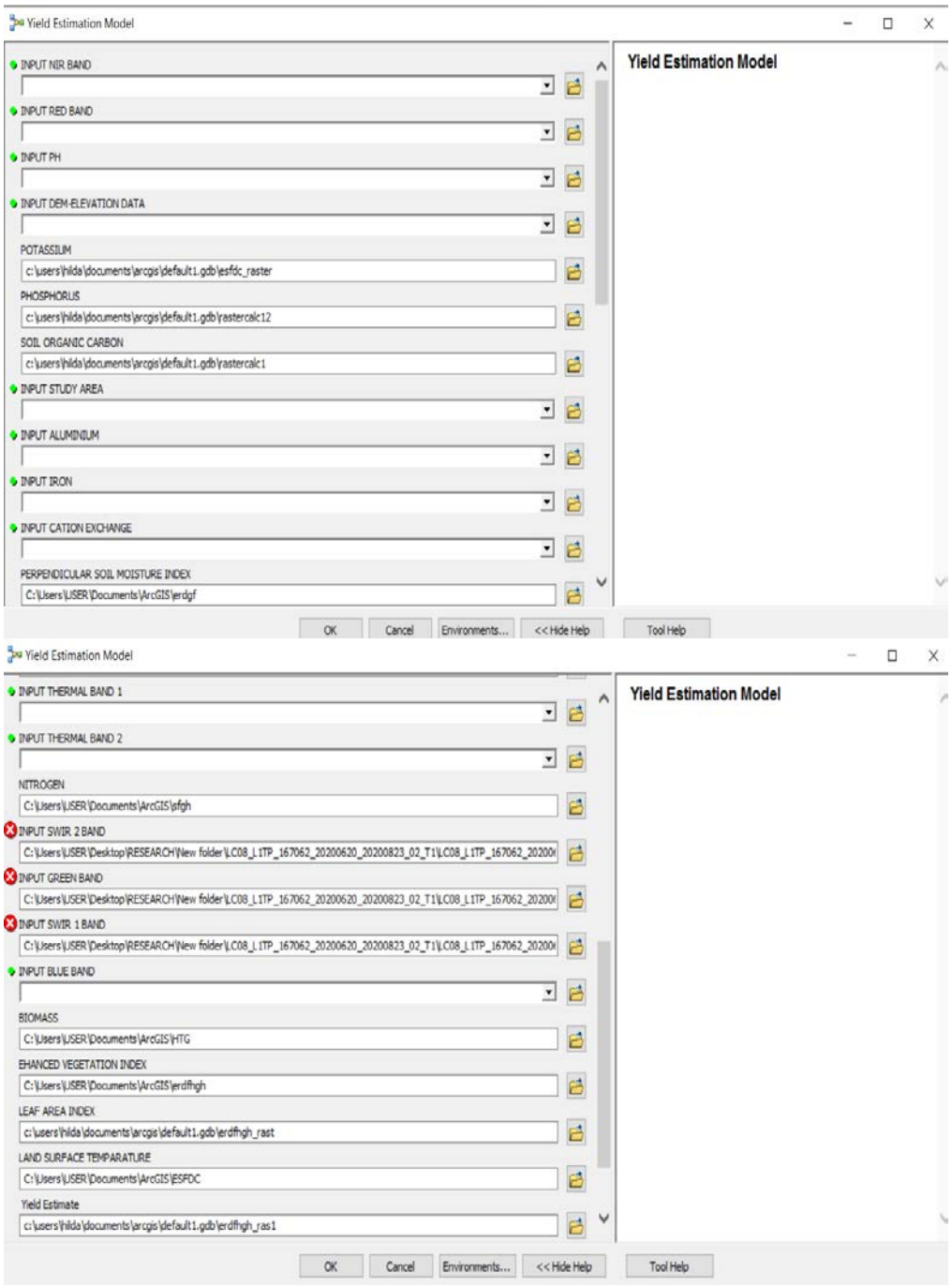


Figure 4.33: Front part of the yield estimation model for green gram and sorghum

4.4.4 Farm data estimation tool

The overall developed farm data estimation tool was converted into Arctool, making its accessibility easy. In Figure 4.59, the FcropDEsti arc tool hosts three models for farm data estimation. The packaged Arctool can be shared and used by anyone who has installed the ARCGIS software. Geographic information systems are tools that enable rational management of spatially distributed resources. This implies that often a GIS enables interaction with either one or many map layers (Liu *et al*., 2011). In the decision-making process, it is common that the base data leading to several rational alternatives is prepared (Nitschke and Innes 2008) at several management positions in an organizational hierarchy. Under these circumstances, a data model developed by a GIS expert following an agreed concept would require repeated computations or repeated decision modifications (Wan *et al*., 2008). Though repeated computations are faster when working with computers, the speed of an activity often depends on the number of tools to be used, the number of working windows, the number of clicks required for a particular operation, the menu selection sequence, the reaction time of a person, etc. (Panda 2003). Therefore, the current study created a GIS database that enables repeated computations without being directly involved in the back-end programming processing for farm area computation, crop identification, agroecological sub-zonation, and yield estimation. The strength of the FCropDesti Tool is to provide or facilitate user-friendliness for people with basic knowledge of GIS to perform farm field data estimation with ease.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

The precision of the crop data estimation under farm field conditions is highly dependent on the stage of the crop, the agroecological zone where most of the climatic characteristics are similar, and the use of one cropping season and one crop growth stage, as was the case for this particular study. In our study, seasonality was found to be crucial as it affected the variability in time series, while regional stratification, for example, the use of agroecological zones, was found to give the best result. The study showed that the classification of fragmented landscapes was improved by using training data and farm boundaries. This aligns with objective one of the research study. A further confirmation of crop identification is the landcover classification carried out in the study area. The TOA reflectance was found to enhance the identification of crops, and the use of multitemporal images was found to be critical in creating spectral signature ranges for the identification of crops. There isn't a single spectral signature through which a certain crop exists, but most crops will tend to oscillate within a certain range for all significant bands (4,5, 6, and 7), as seen in this study. This information can be critically used for the identification of crops and, hence, accurate data estimation. Operating within an agroecological zone, certain crop stages, and a particular season is vital if crop identification is to be successful and data estimation of higher quality. The overall effectiveness of this approach has been demonstrated by the accuracy of the assessment of the landcover classification.

The research study confirms that crop spatial distribution and spectral signature ranges can be determined for different crop types and patterns under farm field conditions. This was achieved through analyzing satellite data from a specific crop stage of growth and including cropping pattern data and farm delineation as part of the methodology for crop data estimation.

Further, through the MCDA analysis, it was possible to confirm that micro-agroecological zones do exist, and data estimation should be restricted to these homogenous zones while working under farm field conditions. This addressed

objective three of the research study. In addition to ensuring proper estimation of the crop area and crop identification, crop yield estimation was of high accuracy. Using machine learning methods such as random forest, 11 parameters that focused on crop health, soil status, and climate status were used. To reach the third goal of the research study, which was to estimate crop areas, we used the spectral signature model along with the micro-agroecological zones model to get very accurate estimates of sorghum and green gram crop yields in farm fields. The three models were aggregated in ARCGIS software to enable data estimation under farm-field conditions. The research provides a stepwise process for earth observation-based crop data estimation under farm field conditions. This aligns with Objective 4 of the research study.

Therefore, it can be concluded that farm delineation is important data for crop data estimation. It enables the identification of cropped areas for spatial distribution. Spectral signature data ranges can be developed for crop identification and crop area estimation. The Top of Atmosphere (TOA) reflectance's play a significant role in the development of the spectral signature ranges. Agro-ecological sub-zonation is very vital in the creation of a homogenous environment in which the parameters for crop yield estimation can be used under farm field conditions. Lastly, a tool for estimating farm data was developed for specific agro-ecological sub-zonation for a certain crop stage on defined farms. It uses spectral signatures to understand crop characteristics and estimate important crop data. An earth observation-based crop data estimation tool under farm field conditions is important and can be developed for smallholder farmers.

The study, however, recommends that the current agro-ecological zones need to undergo micro-climatic zone assessment for improved farm field crop data through the creation of homogenous zones. The research study opens up the possibility of creating spectral crop libraries like those done for the Asian continent for better crop data estimation. The development of spectral libraries for African content should be done to aid in better crop management through proper planning and crop allocation.

Further validation of the modules and the tools is necessary for scalability to other sorghum and green gram-growing regions and application to other crops. There will be a need to validate the optimum performance of the models from this study for crop data estimation for sorghum and green gram. Further testing of the crop data estimation tool is necessary to ensure its stability. Finally, models for different crops can be developed to enable crop data estimation. This study is a contribution to the ongoing efforts towards overcoming challenges faced in remote sensing of agricultural ecosystems, and the resultant formation is potentially crucial for both county and national government policymakers for planning, management, and crop allocation to different ecozones.

REFERENCES

Abd-elmabod, S. K., Bakr, N., Muñoz-rojas, M., and Pereira, P. (2019). *Assessment of Soil Suitability for Improvement of Soil Factors and Agricultural Management*. https://doi.org/10.3390/su11061588

Abdullah-Al-Faisal, Abdulla - Al Kafy, Foyezur Rahman, A. N. M., Rakib, A. Al, Akter, K. S., Raikwar, V., Jahir, D. M. A., Ferdousi, J., and Kona, M. A. (2021). Assessment and prediction of seasonal land surface temperature change using multi-temporal Landsat images and their impacts on agricultural yields in Rajshahi, Bangladesh. *Environmental Challenges*, *4*(May), 100147. https://doi.org/10.1016/j.envc.2021.100147

Adam, M., Corbeels, M., Leffelaar, P. A., Van Keulen, H., Wery, J., and Ewert, F. (2012). Building crop models within different crop modelling frameworks. *Agricultural Systems*, *113*, 57–63. https://doi.org/10.1016/j.agsy.2012.07.010

agra-africa-agriculture-status-report-2014. (n.d.).

AHADY, A. B., and KAPLAN, G. (2021). Classification comparison of Landsat-8 and Sentinel-2 data in Google Earth Engine, study case of the city of Kabul. *International Journal of Engineering and Geosciences*, *7*(1), 24–31. https://doi.org/10.26833/ijeg.860077

Akinci, H., Özalp, A. Y., and Turgut, B. (2013). Agricultural land use suitability analysis using GIS and AHP technique. *Computers and Electronics in Agriculture*, *97*, 71–82. https://doi.org/10.1016/j.compag.2013.07.006

Ali Khan, M., Ahmad, R., and Hasan Khan, H. (2022). Multi-Criteria Land Suitability Analysis for Agriculture Using AHP and Remote Sensing Data of Northern Region India. In *Geographic Information Systems and Applications in Coastal Studies*. IntechOpen. https://doi.org/10.5772/intechopen.102432

Amwata, D. A. (2020). *Final Report SITUATIONAL ANALYSIS OF THE AGRICULTURE SECTOR IN KENYA. July*.

Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Sitch, S., Piao, S., and Zhu, Z. (2013). Evaluation of land surface models in reproducing satellite derived leaf area index over the high-latitude northern hemisphere. Part II: Earth system models. *Remote Sensing*, *5*(8), 3637–3661. https://doi.org/10.3390/rs5083637

Anda, A., and Pinter, L. (1994). Sorghum germination and development as influenced by soil temperature and water content. *Agronomy Journal*, *86*(4), 621–624. https://doi.org/10.2134/agronj1994.00021962008600040008x

Ansarifar, J., Wang, L., and Archontoulis, S. V. (2021). An interaction regression model for crop yield prediction. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-97221-7

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. In *Remote Sensing* (Vol. 5, Issue 2, pp. 949–981). https://doi.org/10.3390/rs5020949

Atzmanstorfer, K., and Blaschke, T. (2013). The Geospatial Web: A Tool to Support the Empowerment of Citizens through E-Participation? *Citizen E-Participation in Urban Governance: Crowdsourcing and Collaborative Creativity*, *July*, 144–170. https://doi.org/10.4018/978-1-4666-4169-3.ch009

Awad, M. M. (2019). Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture (Switzerland)*, *9*(3). https://doi.org/10.3390/agriculture9030054

Awad, M. M., Alawar, B., and Jbeily, R. (2019). A new crop spectral signatures database interactive tool (CSSIT). *Data*, *4*(2), 1–14. https://doi.org/10.3390/data4020077

Ayehu, G. T., and Besufekad, S. A. (2015). Rain fed rice, Land suitability, AHP, MCDM, GIS; Rain fed rice, Land suitability, AHP, MCDM, GIS. *American Journal of Geographic Information System*, *2015*(3), 95–104. https://doi.org/10.5923/j.ajgis.20150403.02

Batjes, N. H., Al-Adamat, R., Bhattacharyya, T., Bernoux, M., Cerri, C. E. P., Gicheru, P., Kamoni, P., Milne, E., Pal, D. K., and Rawajfih, Z. (2007). Preparation of consistent soil data sets for modelling purposes: Secondary SOTER data for four case study areas. *Agriculture, Ecosystems and Environment, 122*(1), 26–34. https://doi.org/10.1016/j.agee.2007.01.005

Berhane, F., and Zaitchik, B. (2014). Modulation of daily precipitation over East Africa by the Madden-Julian oscillation. *Journal of Climate, 27*(15), 6016–6034. https://doi.org/10.1175/JCLI-D-13-00693.1

Birch, I. (2018). Agricultural productivity in Kenya : Barriers and opportunities. *K4D Knowledge, Evidence and Learning for Development, December*, 1–19. http://www.fao.org/faostat

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., and Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sensing of Environment, 269*. https://doi.org/10.1016/j.rse.2021.112831

Bobade, S. V., Bhaskar, B. P., Gaikwad, M. S., Raja, P., Gaikwad, S. S., Anantwar, G., Patil, S. V., Singh, S. R., and Maji, A. K. (2010). A GIS-based land use suitability assessment in Seoni district, Madhya Pradesh, India. *Tropical Ecology*, *51*(1), 41–54.

Bojinski, S., Schaepman, M., Apfer, D. S., and Itten, K. (2003). SPECCHIO: a spectrum database for remote sensing applications. In *Computers and Geosciences* (Vol. 29).

Boote, K. J., Jones, J. W., and Pickering, N. B. (1996). Potential Uses and Limitations of Crop Models I . Model Use as a Research Tool. *Agronomy Journal*, *716*, 704–716.

Bousbih, S., Zribi, M., Pelletier, C., Gorrab, A., Lili-Chabaane, Z., Baghdadi, N., Aissa, N. Ben, and Mougenot, B. (2019). Soil texture estimation using radar and optical data from Sentinel-1 and Sentinel-2. *Remote Sensing*, *11*(13). https://doi.org/10.3390/rs11131520

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiè, F., Cabidoche, Y. M., Cellier, P., Debaeke, P., Gaudillè, J. P., Hénault, C., Maraux, F., Seguin, B., and Sinoquet, H. (2003). *An overview of the crop model STICS*. www.elsevier.com/locate/eja

Bruce, C. M., and Hilbert, D. W. (2004). *Pre-processing Methodology for Application to Landsat TM/ETM+ Imagery of the Wet Tropics*.

Cao, X. L., Zhou, Z. H., Chen, X. D., Shao, W. W., and Wang, Z. R. (2015). Improving leaf area index simulation of IBIS model and its effect on water carbon and energy-A case study in Changbai Mountain broadleaved forest of China. *Ecological Modelling*, *303*, 97–104. https://doi.org/10.1016/j.ecolmodel.2015.02.012

Castillejo-González, I. L., López-Granados, F., García-Ferrer, A., Peña-Barragán, J. M., Jurado-Expósito, M., de la Orden, M. S., and González-Audicana, M. (2009). Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Computers and Electronics in Agriculture*, *68*(2), 207–215. https://doi.org/10.1016/j.compag.2009.06.004

Charles, H., Godfray, J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. (2010). *Food Security: The Challenge of Feeding 9 Billion People*. www.sciencemag.org

Cohen, L., Manion, Lawrence., and Morrison, K. (Keith R. B.). (2007). *Research methods in education*. Routledge.

Cole, S., Killeen, G., Harigaya, T., and Krishna, A. (2023). *Using Satellites and Phones to Evaluate and Promote Agricultural Technology Adoption: Evidence from Smallholder Farms in India **.

Crnojevic, V., Lugonja, P., Brkljac, B., and Brunet, B. (2014). Classification of small agricultural fields using combined Landsat-8 and RapidEye imagery: case study of northern Serbia. *Journal of Applied Remote Sensing*, *8*(1), 083512. https://doi.org/10.1117/1.jrs.8.083512

Cunha, M., Marçal, A. R. S., and Silva, L. (2010). Very early prediction of wine yield based on satellite data from vegetation. *International Journal of Remote Sensing*, *31*(12), 3125–3142. https://doi.org/10.1080/01431160903154382

Cw, M., Rm, M., Kirubi, D. T., Muui, C., Muasya, R., and Kirubi, D. (2013). *BASELINE SURVEY ON FACTORS AFFECTING SORGHUM PRODUCTION AND USE IN EASTERN KENYA* (Vol. 13, Issue 1).

Davidson, D. J. (2019). *EFFECT OF DIFFERENT METHODS AND TIMING OF NITROGEN (N) APPLICATION ON SORGHUM (SORGHUM BICOLOR L) GRAIN YIELD CORE View metadata, citation and similar papers at core.ac.uk provided by SHAREOK repository*.

Dente, L., Satalino, G., Mattia, F., and Rinaldi, M. (2008). Assimilation of leaf area index derived from ASAR and MERIS data into CERES-Wheat model to map wheat yield. *Remote Sensing of Environment*, *112*(4), 1395–1407. https://doi.org/10.1016/j.rse.2007.05.023

Dinh Duong, N., Van Le, A., Nguyen, K.-A., Nguyen Dinh, D., Le Van, A., Ho Le, T., and Kim, Anhn. (2014). *Spectral signatures in landsat 8 oli image and their interpretation for land cover study*. https://www.researchgate.net/publication/287676098

Doraiswamy, P. C., Moulin, S., Cook, P. W., and Stern, A. (2003). *Crop Yield Assessment from Remote Sensing*.

Dorcas, K., Koech, O. K., Kinama, J. M., Chemining'wa, G. N., and Ojulong, H. F. (2019). SORGHUM PRODUCTION PRACTICES IN AN INTEGRATED CROP-LIVESTOCK PRODUCTION SYSTEM IN MAKUENI COUNTY, EASTERN KENYA †. In *Tropical and Subtropical Agroecosystems* (Vol. 22).

Duane Nellis, M., Price, K. P., and Rundquist, D. (2009). *Remote Sensing of Cropland Agriculture*. https://doi.org/10.4135/978-1-8570-2105-9.n26

Ehrgott, M., Figueira, J. R., and Greco, S. (Eds.). (2010). *Trends in Multiple Criteria Decision Analysis* (Vol. 142). Springer US. https://doi.org/10.1007/978-1-4419-5904-1

Eskandari, H., Rabelo, L., and Blvd, F. (n.d.). *HANDLING UNCERTAINTY IN THE ANALYTIC HIERARCHY PROCESS: A STOCHASTIC APPROACH*.

Estévez, J., Salinero-Delgado, M., Berger, K., Pipia, L., Rivera-Caicedo, J. P., Wocher, M., Reyes-Muñoz, P., Tagliabue, G., Boschetti, M., and Verrelst, J. (2022). Gaussian processes retrieval of crop traits in Google Earth Engine based on Sentinel-2 top-of-atmosphere data. *Remote Sensing of Environment*, *273*. https://doi.org/10.1016/j.rse.2022.112958

Everest, T., Sungur, A., and Özcan, H. (2021). Determination of agricultural land suitability with a multiple-criteria decision-making method in Northwestern Turkey. *International Journal of Environmental Science and Technology*, *18*(5), 1073–1088. https://doi.org/10.1007/s13762-020-02869-9

Fageria. (2009). *The Importance of Potassium (K) in Agricultural Soils 10 Things to know about Potassium (K)*.

Faisal, B. M. R., Rahman, H., Sharifee, N. H., Sultana, N., Islam, M. I., Habib, S. M. A., and Ahammad, T. (2020). Integrated Application of Remote Sensing and GIS in Crop Information System—A Case Study on Aman Rice Production Forecasting Using MODIS-NDVI in Bangladesh. *AgriEngineering*, *2*(2), 264–279. https://doi.org/10.3390/agriengineering2020017

Fang, H., and Liang, S. (2003). Retrieving leaf area index with a neural network method: Simulation and validation. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(9 PART I), 2052–2062. https://doi.org/10.1109/TGRS.2003.813493

FAO. (1996). *AGRO-ECOLOGICAL ZONING Guidelines FAO Soils Bulletin 76 Soil Resources, Management and Conservation Service FAO Land and Water Development Division*.

FAO. (2004). *The State of Food and Agriculture: Meeting the needs of the poor?* http://www.fao.org/tempref/docrep/fao/006/y5160e/y5160e.pdf%0Ahttp://www.fao.org/docrep/006/Y5160E/Y5160E00.htm

FAO. (2008). *FOR CLIMATE CHANGE ASSESSMENT, MITIGATION AND ADAPTATION THE GLOBAL TERRESTRIAL OBSERVING SYSTEM*. www.wmo.ch

FAO. (2011). The state of food and agriculture: Women in agriculture (executive summary). *Lancet*, *2*(7929), 160. http://www.ncbi.nlm.nih.gov/pubmed/24897208%5Cnhttp://www.fao.org/catalog/inter-e.htm

FAO. (2013). *Crop Yield Forecasting: Methodological and Institutional Aspects*. 241.

FAO. (2014). *Bioenergy and Food Security User Manual Crop Production*. 1–36. papers2://publication/uuid/870B0622-8202-4CFF-963A-08C24F52A7D3

Ferencz, C., Bognár, P., Lichtenberger, J., Hamar, D., Tarcsai, G., Timár, G., Molnár, G., Pásztor, S., Steinbach, P., Székely, B., Ferencz, O. E., and Ferencz-Árkos, I. (2004). Crop yield estimation by satellite remote sensing. *International Journal of Remote Sensing*, *25*(20), 4113–4149. https://doi.org/10.1080/01431160410001698870

Fernandes, R., Plummer, S., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M., Lacaze, R., Leblanc, S., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J., … Nickeson, J. (2014). *Committee on Earth Observation Satellites Working Group on Calibration and Validation Land Product Validation Sub-Group Global Leaf Area Index Product Validation Good Practices Editors*.

Feyisa, G. L., Palao, L. K., Nelson, A., Gumma, M. K., Paliwal, A., Win, K. T., Nge, K. H., and Johnson, D. E. (2020). Characterizing and mapping cropping patterns in a complex agro-ecosystem: An iterative participatory mapping procedure using machine learning algorithms and MODIS vegetation indices. *Computers and Electronics in Agriculture*, *175*. https://doi.org/10.1016/j.compag.2020.105595

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., … Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342. https://doi.org/10.1038/nature10452

Forkuor, G., Dimobe, K., Serme, I., and Tondoh, J. E. (2018). Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience and Remote Sensing*, *55*(3), 331–354. https://doi.org/10.1080/15481603.2017.1370169

Fox, D. M., Bryan, R. B., and Price, A. G. (1997). The influence of slope angle on final infiltration rate for interrill conditions. In *Geoderma* (Vol. 80).

Fritz, S., Massart, M., Savin, I., Gallego, J., and Rembold, F. (2008). The use of MODIS data to derive acreage estimations for larger fields: A case study in the south-western Rostov region of Russia. *International Journal of Applied Earth Observation and Geoinformation*, *10*(4), 453–466. https://doi.org/10.1016/j.jag.2007.12.004

Fritz, S., See, L., Bayas, J. C. L., Waldner, F., Jacques, D., Becker-Reshef, I., Whitcraft, A., Baruth, B., Bonifacio, R., Crutchfield, J., Rembold, F., Rojas, O., Schucknecht, A., Van der Velde, M., Verdin, J., Wu, B., Yan, N., You, L., Gilliams, S., … McCallum, I. (2019). A comparison of global agricultural monitoring systems and current gaps. *Agricultural Systems*, *168*, 258–272. https://doi.org/10.1016/j.agsy.2018.05.010

Gao, F., and Zhang, X. (2021). Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities. *Journal of Remote Sensing*, *2021*, 1–14. https://doi.org/10.34133/2021/8379391

García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., and Vázquez-Peña, M. A. (2020). Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and rgb images acquired with unmanned aerial vehicles. *Agriculture (Switzerland)*, *10*(7), 1–24. https://doi.org/10.3390/agriculture10070277

Garriga, M., Retamales, J. B., Romero-Bravo, S., Caligari, P. D. S., and Lobos, G. A. (2014). Chlorophyll, anthocyanin, and gas exchange changes assessed by spectroradiometry in Fragaria chiloensis under salt stress. *Journal of Integrative Plant Biology*, *56*(5), 505–515. https://doi.org/10.1111/jipb.12193

Ghazali, M. F., Wikantika, K., Harto, A. B., and Kondoh, A. (2020). Generating soil salinity, soil moisture, soil pH from satellite imagery and its analysis. *Information Processing in Agriculture*, *7*(2), 294–306. https://doi.org/10.1016/j.inpa.2019.08.003

Graesser, J., and Ramankutty, N. (2017). Detection of cropland field parcels from Landsat imagery. *Remote Sensing of Environment*, *201*, 165–180. https://doi.org/10.1016/j.rse.2017.08.027

Hansen, J. W., and Indeje, M. (2004). Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agricultural and Forest Meteorology*, *125*(1–2), 143–157. https://doi.org/10.1016/j.agrformet.2004.02.006

HAO, P. yu, TANG, H. jun, CHEN, Z. xin, MENG, Q. yan, and KANG, Y. peng. (2020). Early-season crop type mapping using 30-m reference time series. *Journal of Integrative Agriculture*, *19*(7), 1897–1911. https://doi.org/10.1016/S2095-3119(19)62812-1

Heupel, K., Spengler, D., and Itzerott, S. (2018). A Progressive Crop-Type Classification Using Multitemporal Remote Sensing Data and Phenological Information. *PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, *86*(2), 53–69. https://doi.org/10.1007/s41064-018-0050-7

Hills, R. C. (1979). The structure of the inter-tropical convergence zone in equatorial Africa and its relationship to east African rainfall. *Transactions Institute of British Geographers*, *4*(3), 329–352. https://doi.org/10.2307/622055

Hoefsloot, P., Ines, A., Dam, J. Van, Duveiller, G., Kayitakire, F., and Hansen, J. (2012). *No Title*.

Hoogenboom, G., Porter, C. H., Boote, K. J., Shelia, V., Wilkens, P. W., Singh, U., White, J. W., Asseng, S., Lizaso, J. I., Moreno, L. P., Pavan, W., Ogoshi, R., Hunt, L. A., Tsuji, G. Y., and Jones, J. W. (2019). *The DSSAT crop modeling ecosystem* (pp. 173–216). https://doi.org/10.19103/as.2019.0061.10

Hossen, B., Yabar, H., and Mizunoya, T. (2021). Land suitability assessment for pulse (Green gram) production through remote sensing, GIS and multicriteria analysis in the coastal region of Bangladesh. *Sustainability (Switzerland)*, *13*(22). https://doi.org/10.3390/su132212360

Huang, X., Liu, J., Zhu, W., Atzberger, C., and Liu, Q. (2019). The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232725

Hueni, A., Nieke, J., Schopfer, J., Kneubühler, M., and Itten, K. I. (2009). The spectral database SPECCHIO for improved long-term usability and data sharing. *Computers and Geosciences*, *35*(3), 557–565. https://doi.org/10.1016/j.cageo.2008.03.015

Hui, J., and Yao, L. (2018). A Method to Upscale the Leaf Area Index (LAI) Using GF-1 Data with the Assistance of MODIS Products in the Poyang Lake Watershed. *Journal of the Indian Society of Remote Sensing*, *46*(4), 551–560. https://doi.org/10.1007/s12524-017-0731-5

Husak, G. J., Marshall, M. T., Michaelsen, J., Pedreros, D., Funk, C., and Galu, G. (2008). Crop area estimation using high and medium resolution satellite imagery in areas with complex topography. *Journal of Geophysical Research Atmospheres*, *113*(14). https://doi.org/10.1029/2007JD009175

Idso, S. B., Jackson, R. D., and Reginato, R. J. (1977). REMOTE SENSING FOR AGRICULTURAL WATER MANAGEMENT AND CROP YIELD PREDICTION. In *Agricultural Water Management* (Vol. 1).

Iodice, F., D'Acunto, F., and Bigagli, L. (2021). Mapping nitrogen from satellite data to improve soil quality - A worked example. *GI_Forum*, *9*(1), 104–111. https://doi.org/10.1553/GISCIENCE2021_01_S104

Ishimwe, R. (2014). *INDENTIFYING CROPS USING LANDSAT 8 THERMAL INFRARED BANDS*.

J Bolt, C Duku, A Groot, T Demissie, J. R., and Bolt, J., Duku, C., Groot, A., T Demissie, T., Recha, J. (2019). *Green Grams Kenya: Climate Change Risks and Opportunities*. 1–4.

Jamal-Eddine, O., Harti Abderrazak, E., Rachid, L., Moujahid Ali, E., Naima, B., Ouazzani Rabii, E., El Mostafa, B., Ghmari Abderrahmene, E., and Mellal, B. (2018). *Crop type mapping from pansharpened Landsat 8 NDVI data: A case of a highly fragmented and intensive agricultural system*.

Jat, M. K., Yadav, P. K., Tikkoo, A., Ss, and, Chaudhary, Y., and Singh, C. (2022). Effect of potassium application on yield and quality of green gram (Vigna radiata L.) on coarse textured soils of southern Haryana. In *Indian Journal of Experimental Biology* (Vol. 60).

Jensen, J. R. (2000). *Remote sensing of the environment : an earth resource perspective*. Pearson Prentice Hall.

Jiitzold, R., and Kutsch', H. (2000). *Der Tropenlandwirt, Zeitschrift fur die Landwirtschaft in den Tropen und Subtropen Agro-Ecological Zones of the Tropics, with a Sample from Kenya Calculation of decadic values for kc by a new formula set according to H. Kutsch*.

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., and Wheeler, T. R. (2017). Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agricultural Systems*, *155*, 269–288. https://doi.org/10.1016/j.agsy.2016.09.021

K. Boitt, M., N. Mundia, harles, and Pellikka, P. (2014). Modelling the Impacts of Climate Change on Agro-Ecological Zones – a Case Study of Taita Hills, Kenya. *Universal Journal of Geoscience*, *2*(6), 172–179. https://doi.org/10.13189/ujg.2014.020602

Kachhwaha, T. S. (1983). Spectral signatures obtained from Landsat digital data for forest vegetation and land-use mapping in India. *Photogrammetric Engineering and Remote Sensing*, *49*(5), 685–689.

Kahsay, A., Haile, M., Gebresamuel, G., Mohammed, M., and Moral, M. T. (2018). Land suitability analysis for sorghum crop production in northern semi-arid Ethiopia: Application of GIS-based fuzzy AHP approach. *Cogent Food and Agriculture*, *4*(1), 1–24. https://doi.org/10.1080/23311932.2018.1507184

Kamara, S. I., and Jackson, I. J. (1997). A new soil-moisture based classification of raindays and drydays and its application to Sierra Leone. *Theoretical and Applied Climatology*, *56*(3–4), 199–213. https://doi.org/10.1007/BF00866427

Kang, Y., Özdoğan, M., Zipper, S. C., Román, M. O., Walker, J., Hong, S. Y., Marshall, M., Magliulo, V., Moreno, J., Alonso, L., Miyata, A., Kimball, B., and Loheide, S. P. (2016). How universal is the relationship between remotely sensed vegetation indices and crop leaf area index? A global assessment. *Remote Sensing*, *8*(7). https://doi.org/10.3390/rs8070597

Karienye, J. M., Nduru, G. M., Huho, J. M., and Opiyo, F. E. (2019). *Small Scale Farmers in Kieni East Sub County , Kenya*. *08*(02), 7–19.

Keating, B. A., Herrero, M., Carberry, P. S., Gardner, J., and Cole, M. B. (2014). Food wedges: Framing the global food demand and supply challenge towards 2050. In *Global Food Security* (Vol. 3, Issues 3–4, pp. 125–132). Elsevier B.V. https://doi.org/10.1016/j.gfs.2014.08.004

Kigen, C. K., Ochieno, M. W., Muoma, J. O., Shivoga, W. A., Konje, M., Onyando, Z., Soi, B. C., Makindi, S. M., Kisoyan, P. M., and Mironga, J. M. (2014). *Spatial Modeling of Sorghum (Sorghum bicolor) growing areas in Kenyan Arid and Semi-Arid Lands*. *66*, 20674–20678.

Kihara, J., Bolo, P., Kinyua, M., Rurinda, J and Piikki, K. (2020). Micronutrient deficiencies in African soils and the human nutritional nexus: opportunities with staple crops. Environ Geochem Health 42:3015–3033

Kim, J., Grunwald, S., and Rivero, R. G. (2014). Soil phosphorus and nitrogen predictions across spatial escalating scales in an aquatic ecosystem using remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(10), 6724–6737. https://doi.org/10.1109/TGRS.2014.2301443

Kogan, F., Salazar, L., and Roytman, L. (2012). Forecasting crop production using satellite-based vegetation health indices in Kansas, USA. *International Journal of Remote Sensing*, *33*(9), 2798–2814. https://doi.org/10.1080/01431161.2011.621464

Koomson, E., Muoni, T., Marohn, C., Nziguheba, G., Öborn, I., and Cadisch, G. (2020). Critical slope length for soil loss mitigation in maize-bean cropping systems in SW Kenya. *Geoderma Regional*, *22*, e00311. https://doi.org/10.1016/j.geodrs.2020.e00311

Kr Maurya, A. (2011). *Estimation of Acreage and Crop Production through Remote Sensing and GIS Technique National initiative on Climate Resilient Agriculture (NICRA) View project National initiative on Climate Resilient Agriculture (NICRA) View project*. https://www.researchgate.net/publication/303880320

Krishnan R, Ramachandran R, Murali Mohan ASRKV, Radhadevi P V, P. S. K. and C. R. (1998). *Satellite Data Preprocessing - New Perspectives*. 90–98. http://www.isprs.org/proceedings/XXXII/part1/90_XXXII-part1.pdf

Kumar, P., Narendra Kumar, S., and Shukla, S. (2022). ACREAGE ESTIMATION OF SUGARCANE CROP IN SITAPUR DISTRICT, UTTAR PRADESH USING MICROWAVE SENTINEL DATA. *International Research Journal of Engineering and Technology*. www.irjet.net

Kyei-Mensah, C., Kyerematen, R., and Adu-Acheampong, S. (2019). Impact of Rainfall Variability on Crop Production within the Worobong Ecological Area of Fanteakwa District, Ghana. *Advances in Agriculture*, *2019*. https://doi.org/10.1155/2019/7930127

Lamsal, A., Welch, S. M., Jones, J. W., Boote, K. J., Asebedo, A., Crain, J., Wang, X., Boyer, W., Giri, A., Frink, E., Xu, X., Gundy, G., Ou, J., and Arachchige, P. G. (2017). Efficient crop model parameter estimation and site characterization using large breeding trial data sets. *Agricultural Systems*, *157*, 170–184. https://doi.org/10.1016/j.agsy.2017.07.016

Lang, P., Zhang, L., Huang, C., Chen, J., Kang, X., Zhang, Z., and Tong, Q. (2023). Integrating environmental and satellite data to estimate county-level cotton yield in Xinjiang Province. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.1048479

Launay, M., and Guerif, M. (2005). Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. *Agriculture, Ecosystems and Environment*, *111*(1–4), 321–339. https://doi.org/10.1016/j.agee.2005.06.005

Laurent, V. C. E., Schaepman, M. E., Verhoef, W., Weyermann, J., and Chávez, R. O. (2014). Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance image. *Remote Sensing of Environment*, *140*, 318–329. https://doi.org/10.1016/j.rse.2013.09.005

Li, D., Miao, Y., Gupta, S. K., Rosen, C. J., Yuan, F., Wang, C., Wang, L., and Huang, Y. (2021). Improving potato yield prediction by combining cultivar information and uav remote sensing data using machine learning. *Remote Sensing*, *13*(16). https://doi.org/10.3390/rs13163322

Li, L., Friedl, M. A., Xin, Q., Gray, J., Pan, Y., and Frolking, S. (2014). Mapping crop cycles in China using MODIS-EVI time series. *Remote Sensing*, *6*(3), 2473–2493. https://doi.org/10.3390/rs6032473

Liang, S., Fang, H., and Chen, M. (2001). Atmospheric correction of Landsat ETM+ land surface imagery-Part I: Methods. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(11), 2490–2498. https://doi.org/10.1109/36.964986

Liang, S., Schaepman, M., and Kneubühler, M. (2008). Remote sensing signatures: Measurements, modelling and applications. *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*, 127–143. https://doi.org/10.1201/9780203888445-17

Liu, H. Q., and Huete, A. (1995). Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(2), 457–465. https://doi.org/10.1109/36.377946

Liu, Q., Liang, S., Xiao, Z., and Fang, H. (2014). Retrieval of leaf area index using temporal, spectral, and angular information from multiple satellite data. *Remote Sensing of Environment, 145*, 25–37. https://doi.org/10.1016/j.rse.2014.01.021

Lobell, D. B., and Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology*, *160*(4), 1686–1697. https://doi.org/10.1104/pp.112.208298

Luwesi, C. N., Obando, J. A., and Shisanya, C. A. (2017). The impact of a warming micro-climate on muooni farmers of Kenya. *Agriculture (Switzerland)*, *7*(3). https://doi.org/10.3390/agriculture7030020

Makowski, D., Hillier, J., Wallach, D., Andrieu, B., and Jeuffroy, M.-H. (2006). Parameter Estimation for Crop Models. *Working with Dynamic Crop Models, May 2014*, 55–100.

Manzi K H, Ngene S and Gweyi-Onyango J P (2023 a). Use of Satellite Data to Extract pH Values for Crop Planning and Management of Sorghum and Green Gram in Tharaka Nithi and Machakos Counties, Kenya. Asian Journal of Advances in Agricultural Research. 23 (3): 25-32. DOI: 10.9734/AJAAR/2023/v23i3464

Manzi K H, Ngene S and Gweyi-Onyango J P (2023 b). Use of Satellite-based Leaf Area Index Data for Monitoring Green Gram Crop Growth in Ikombe-Katangi Area, Machakos County, Kenya. Asian Journal of Advances in Agricultural Research. 23 (3): 53-63. DOI: 10.9734/AJAAR/2023/v23i346

Manzi K.H and Gweyi-Onyango J P (2022). Geoinformation for Land Suitability Modelling for Climate-Smart Farming in Africa. Springer Nature Switzerland AG 2022. A. Kumar et al. (eds.), Agriculture, Livestock Production and Aquaculture, https://doi.org/10.1007/978-3-030-93258-9_9: pp 155-165

Manzi, H., and Gweyi-Onyango, J. P. (2020). Agro-ecological Lower Midland Zones IV and V in Kenya Using GIS and Remote Sensing for Climate-Smart Crop Management. In *African Handbook of Climate Change Adaptation* (pp. 1–27). Springer International Publishing. https://doi.org/10.1007/978-3-030-42091-8_35-1

Marshall, M. T., Funk, C., Michaelsen, J., Marshall, M. T. ;, and Funk, C. ; (2012). *Agricultural Drought Monitoring in Kenya Using Evapotranspiration Derived from Remote Sensing and Reanalysis Data*. http://digitalcommons.unl.edu/usgsstaffpubhttp://digitalcommons.unl.edu/usgsstaffpub/978

Mashaba-Munghemezulu, Z., Chirima, G. J., and Munghemezulu, C. (2021). Modeling the spatial distribution of soil nitrogen content at smallholder maize farms using machine learning regression and sentinel-2 data. *Sustainability (Switzerland)*, *13*(21). https://doi.org/10.3390/su132111591

Mashao, J., and Prinsloo, T. (1994). Sorghum production. *Sorghum Production Guide*, 1–14.

Matvienko, I., Gasanov, M., Petrovskaia, A., Jana, R. B., Pukalchik, M., and Oseledets, I. (2020). *Bayesian aggregation improves traditional single image crop classification approaches*. http://arxiv.org/abs/2004.03468

Meehl, and Stocker. (n.d.). *IPCC WG1, 0-Draft: Chapter 10*.

Mishra, A., Hansen, J. W., Dingkuhn, M., Baron, C., Ndiaye, O., Ward, M. N., and Traore, S. B. (2008). *Sorghum yield prediction from seasonal rainfall forecasts in Burkina Faso*. *148*, 1798–1814. https://doi.org/10.1016/j.agrformet.2008.06.007

Mistri, P., and Sengupta, S. (2020). Multi-criteria Decision-Making Approaches to Agricultural Land Suitability Classification of Malda District, Eastern India. *Natural Resources Research*, *29*(3), 2237–2256. https://doi.org/10.1007/s11053-019-09556-8

Mitran, T., Meena, R. S., Lal, R., Layek, J., Kumar, S., and Datta, R. (2018). Role of Soil Phosphorus on Legume Production. In *Legumes for Soil Health and Sustainable Management* (pp. 487–510). Springer Singapore. https://doi.org/10.1007/978-981-13-0253-4_15

Mohamed, E. S., Ali, A., El-Shirbeny, M., Abutaleb, K., and Shaddad, S. M. (2020). Mapping soil moisture and their correlation with crop pattern using remotely sensed data in arid region. *Egyptian Journal of Remote Sensing and Space Science*, *23*(3), 347–353. https://doi.org/10.1016/j.ejrs.2019.04.003

Moulin, S., Bondeau, A., and Delecolle, R. (1998). Combining agricultural crop models and satellite observations: From field to regional scales. In *International Journal of Remote Sensing* (Vol. 19, Issue 6, pp. 1021–1036). https://doi.org/10.1080/014311698215586

Mtibaa, S., and Irie, M. (2016). Land cover mapping in cropland dominated area using information on vegetation phenology and multi-seasonal Landsat 8 images. *Euro-Mediterranean Journal for Environmental Integration*, *1*(1), 1–16.

Mudrik Masoud, K. (2019). Detection Of Agricultural Field Boundaries From Sentinel-2 Images Using Fully Convolutional Networks and Super Resolution Mapping.

Mugo, J. N., Karanja, N. N., Gachene, C. K., Dittert, K., Nyawade, S. O., and Schulte-Geldermann, E. (2020). Assessment of soil fertility and potato crop nutrient status in central and eastern highlands of Kenya. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-64036-x

Mugo, J. W., Kariuki, P. C., and Musembi, D. K. (2016). Identification of Suitable Land for Green Gram Production Using GIS Based Analytical Hierarchy Process in Kitui County, Kenya. *Journal of Remote Sensing and GIS*, *5*(3). https://doi.org/10.4172/2469-4134.1000170

Murphy, J., Casley, D. J., and Curry, J. J. (1991). *Farmers' Estimations as a Source of Production Data-Methodological Guidelines for Cereals in Africa*.

Mustafa, A. A., Singh, M., Sahoo, R. N., Ahmed, N., Khanna, M., Sarangi, A., and Mishra, A. K. (2011). *Land Suitability Analysis for Different Crops: A Multi Criteria Decision Making Approach using Remote Sensing and GIS*. http://www.sciencepub.net/researcher

Muzira, N. M., Mushore, T. D., Wuta, M., Mutasa, C., and Mashonjowa, E. (2021). Land suitability analysis of Zimbabwe for the production of sorghum (Sorghum -bicolor) and maize (Zea mays) using a Remote Sensing and GIS based approach. *Remote Sensing Applications: Society and Environment*, *23*. https://doi.org/10.1016/j.rsase.2021.100553

Mwendwa, S. M., Mbuvi, J. P., Kironchi, G., and Gachene, C. K. K. (2022). Assessing spatial variability of selected soil properties in Upper Kabete Campus coffee farm, University of Nairobi, Kenya. *Heliyon*, *8*(8). https://doi.org/10.1016/j.heliyon.2022.e10190

Myneni, R. B., and Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*, *49*(3), 200–211. https://doi.org/10.1016/0034-4257(94)90016-7

Natarajan, V. A., Kumar, M. S., Tamizhazhagan, V., and Chevdumoi, R. M. (2022). PREDICTION OF SOIL PH FROM REMOTE SENSING DATA USING GRADIENT BOOSTED REGRESSION ANALYSIS. *Journal of Pharmaceutical Negative Results ¦*, *13*, 2022. https://doi.org/10.47750/pnr.2022.13.S06.005

Nawaz, A., Iqbal, Z., and Ullah, S. (2010). *Performance Analysis of Supervised Image Classification Techniques for the Classification of Multispectral Satellite Imagery*.

Nellis, M. D., Price, K. P., and Rundquist, D. (2008). Remote Sensing of Cropland Agriculture. *The SAGE Handbook of Remote Sensing*, 368–383. https://doi.org/10.4135/9780857021052.n26

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., and McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, *352*(May), 251–267. https://doi.org/10.1016/j.geoderma.2019.06.016

Ngure, M. W., Wandiga, S. O., Olago, D. O., and Oriaso, S. O. (2021). Climate change stressors affecting household food security among Kimandi-Wanyaga smallholder farmers in Murang'a County, Kenya. *Open Agriculture*, *6*(1), 587–608. https://doi.org/10.1515/opag-2021-0042

Nguy-Robertson, A., Gitelson, A., Peng, Y., Walter-Shea, E., Leavitt, B., and Arkebauer, T. (2013). Continuous monitoring of crop reflectance, vegetation fraction, and identification of developmental stages using a four band radiometer. *Agronomy Journal*, *105*(6), 1769–1779. https://doi.org/10.2134/agronj2013.0242

Nidamanuri, R. R., and Zbell, B. (2012). Existence of characteristic spectral signatures for agricultural crops - Potential for automated crop mapping by hyperspectral imaging. *Geocarto International*, *27*(2), 103–118. https://doi.org/10.1080/10106049.2011.623792

Onduru, D. D., and Preez, C. C. Du. (2008). *Spatial and temporal aspects of agricultural sustainability in the semi-arid tropics : a case study in Mbeere district , Eastern Kenya*. *47*(January), 134–148.

Orr, A., Mwema, C., and Mulinge, W. (2013). *Socioeconomics Discussion Paper Series The value chain for sorghum beer in Kenya*. www.icrisat.org

Padma, S., and Sanjeevi, S. (2014). Jeffries Matusita based mixed-measure for improved spectralmatching in hyperspectral image analysis. *International Journal of Applied Earth Observation and Geoinformation*, *32*(1), 138–151. https://doi.org/10.1016/j.jag.2014.04.001

Panda, S. S., Ames, D. P., and Panigrahi, S. (2010). Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sensing*, *2*(3), 673–696. https://doi.org/10.3390/rs2030673

Panigrahy, S., Ray, S. S., Manjunath, K. R., Pandey, P. S., Sharma, S. K., Sood, A., Yadav, M., Gupta, P. C., Kundu, N., and Parihar, J. S. (2011). A Spatial Database of Cropping System and its Characteristics to Aid Climate Change Impact Assessment Studies. *Journal of the Indian Society of Remote Sensing*, *39*(3), 355–364. https://doi.org/10.1007/s12524-011-0093-3

Parker, G. G. (2020). Tamm review: Leaf Area Index (LAI) is both a determinant and a consequence of important processes in vegetation canopies. *Forest Ecology and Management*, *477*(June). https://doi.org/10.1016/j.foreco.2020.118496

Pellarin, T., Román-Cascón, C., Baron, C., Bindlish, R., Brocca, L., Camberlin, P., Fernández-Prieto, D., Kerr, Y. H., Massari, C., Panthou, G., Perrimond, B., Philippon, N., and Quantin, G. (2020). The precipitation inferred from soil moisture (PrISM) near real-time rainfall product: Evaluation and comparison. *Remote Sensing*, *12*(3). https://doi.org/10.3390/rs12030481

Potgieter, A. B., Zhao, Y., Zarco-Tejada, P. J., Chenu, K., Zhang, Y., Porker, K., Biddulph, B., Dang, Y. P., Neale, T., Roosta, F., and Chapman, S. (2021). Evolution and application of digital technologies to predict crop type and crop phenology in agriculture. In *In Silico Plants* (Vol. 3, Issue 1). Oxford University Press. https://doi.org/10.1093/insilicoplants/diab017

Pratap Singh, D. (2017). Effect of potassium and sulphur on performance of green gram (Vigna radiata) in alluvial soil. In *Annals of Plant and Soil Research* (Vol. 19, Issue 2).

Qiu, B., Hu, X., Chen, C., Tang, Z., Yang, P., Zhu, X., Yan, C., and Jian, Z. (2022). Maps of cropping patterns in China during 2015–2021. *Scientific Data*, *9*(1). https://doi.org/10.1038/s41597-022-01589-8

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A. (2008). Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, *22*(1). https://doi.org/10.1029/2007GB002952

Ray, S. S., Singh, J. P., Das, G., and Panigrahy, S. (2002). *USE OF HIGH RESOLUTION REMOTE SENSING DATA FOR GENERATING SITE-SPECIFIC SOIL MANGEMENT PLAN*. http://www.spaceimaging.com/products/ikonos/spectral.html

Region, C. M. (2018). *Soil Suitability Analysis for Sustainable Landuse Planning in Maheshkhola Soil Suitability Analysis for Sustainable Landuse Planning in Maheshkhola Watershed , Central Mountain Region , Nepal*. *January 2013*.

Regulatory framework for agricultural data in the Near East and North Africa region. (2023). https://doi.org/10.4060/cc6871en

Rembold, F., Atzberger, C., Savin, I., and Rojas, O. (2013). Using low resolution satellite imagery for yield prediction and yield anomaly detection. In *Remote Sensing* (Vol. 5, Issue 4, pp. 1704–1733). https://doi.org/10.3390/rs5041704

Ren, X., Xu, W., and Smith, A. (2012). *Send Orders of Reprints at Reprints@benthamscience.org 68 The Open Hydrology Journal* (Vol. 6).

Reynolds, C. A., Yitayew, M., Slack, D. C., Hutchinson, C. F., Huete, A., and Petersen, M. S. (2000). Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data. *International Journal of Remote Sensing*, *21*(18), 3487–3508. https://doi.org/10.1080/014311600750037516

Richards, J. A. (n.d.). *Remote Sensing Digital Image Analysis*.

Rossato, L., Alvalá, R. C. dos S., Marengo, J. A., Zeri, M., Cunha, A. P. M. do A., Pires, L. B. M., and Barbosa, H. A. (2017). Impact of soil moisture on crop yields over Brazilian semiarid. *Frontiers in Environmental Science*, *5*(NOV). https://doi.org/10.3389/fenvs.2017.00073

Roy, D. P., Zhang, H. K., Ju, J., Gomez-Dans, J. L., Lewis, P. E., Schaaf, C. B., Sun, Q., Li, J., Huang, H., and Kovalskyy, V. (2016). A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sensing of Environment, 176*, 255–271. https://doi.org/10.1016/j.rse.2016.01.023

Rumora, L., Miler, M., and Medak, D. (2021). Contemporary comparative assessment of atmospheric correction influence on radiometric indices between Sentinel-2A and Landsat 8 imagery. *Geocarto International, 36*(1), 13–27. https://doi.org/10.1080/10106049.2019.1590465

Saatsaz, M., Monsef, I., Rahmani, M., and Ghods, A. (2018). Site suitability evaluation of an old operating landfill using AHP and GIS techniques and integrated hydrogeological and geophysical surveys. *Environmental Monitoring and Assessment*, *190*(3). https://doi.org/10.1007/s10661-018-6505-x

Saini, R., and Ghosh, S. K. (2018). CROP CLASSIFICATION ON SINGLE DATE SENTINEL-2 IMAGERY USING RANDOM FOREST AND SUPPOR VECTOR MACHINE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–5*, 683–688. https://doi.org/10.5194/isprs-archives-xlii-5-683-2018

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., and Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, *96*(3–4), 366–374. https://doi.org/10.1016/j.rse.2005.03.008

Samwel, M. P., Abila, R., and Mabwoga, S. (2021). Assessment of Climate Variability in Kisii Kenya and Its Implications on Food Security. *American Journal of Climate Change*, *10*(04), 386–395. https://doi.org/10.4236/ajcc.2021.104019

Sapkota, T. B., Jat, M. L., Aryal, J. P., Jat, R. K., and Khatri-Chhetri, A. (2015). Climate change adaptation, greenhouse gas mitigation and economic profitability of conservation agriculture: Some examples from cereal systems of Indo-Gangetic Plains. *Journal of Integrative Agriculture*, *14*(8), 1524–1533. https://doi.org/10.1016/S2095-3119(15)61093-0

Sapkota, T. B., Jat, M. L., Jat, R. K., Kapoor, P., and Stirling, C. (2016). Yield estimation of food and non-food crops in smallholder production systems. In *Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture* (pp. 163–174). Springer International Publishing. https://doi.org/10.1007/978-3-319-29794-1_8

Selim, S., Koc-San, D., Selim, C., and San, B. T. (2018). Site selection for avocado cultivation using GIS and multi-criteria decision analyses: Case study of Antalya, Turkey. *Computers and Electronics in Agriculture*, *154*, 450–459. https://doi.org/10.1016/j.compag.2018.09.038

Service, U. R. C., and Moines, D. (2018). Estimating Cover Crop Biomass. *USDA-Natural Resources Conservation Service*. https://www.nrcs.usda.gov/wps/cmis_proxy/https/ecm.nrcs.usda.gov%3A443/fncmis/resources/WEBP/ContentStream/idd_F025D465-0000-C114-83EC-729157961659/0/EstBiomassCoverCrops_Sept2018.pdf

Shafian, S., and Maas, S. J. (2015). Index of soil moisture using raw Landsat image digital count data in Texas High Plains. *Remote Sensing*, *7*(3), 2352–2372. https://doi.org/10.3390/rs70302352

Shammi, S. A., and Meng, Q. (2021). Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. *Ecological Indicators*, *121*(June 2020), 107124. https://doi.org/10.1016/j.ecolind.2020.107124

Shelestov, A., Lavreniuk, M., Kussul, N., and Novikov, A. (2017). *Exploring Google Earth Engine Platform for Big Data Processing : Classification of Multi-Temporal Satellite Imagery for Crop Mapping*. *5*(February), 1–10. https://doi.org/10.3389/feart.2017.00017

Shen, M., Chen, J., Zhu, X., and Tang, Y. (2009). *Yellow flowers can decrease NDVI and EVI values: evidence from a field experiment in an alpine meadow*. http://pubs.nrc-cnrc.gc.ca/cjrs

Shi, X., Deng, Z., Ding, X., and Li, L. (2020). Land cover classification combining Sentinel-1 and Landsat 8 imagery driven by Markov random field with amendment reliability factors. *Eurasip Journal on Wireless Communications and Networking*, *2020*(1). https://doi.org/10.1186/s13638-020-01713-5

Shoko, C., Mutanga, O., and Dube, T. (2018). Determining optimal new generation satellite derived metrics for accurate C3 and C4 grass species aboveground biomass estimation in South Africa. *Remote Sensing*, *10*(4). https://doi.org/10.3390/rs10040564

Sibanda, L. M., and Mwamakamba, S. N. (2021). Policy considerations for african food systems: Towards the united nations 2021 food systems summit. *Sustainability (Switzerland)*, *13*(16), 1–15. https://doi.org/10.3390/su13169018

Singh, R., Semwal, D. P., Rai, A., and Chhikara, R. S. (2002). Small area estimation of crop yield using remote sensing satellitie data. *International Journal of Remote Sensing*, *23*(1), 49–56. https://doi.org/10.1080/01431160010014756

Singha, C., and Swain, K. C. (2016). Land suitability evaluation criteria for agricultural crop selection: A review. *Agricultural Reviews*, *37*(2), 125–132. https://doi.org/10.18805/ar.v37i2.10737

Sobrino, J. A., and Jiménez-Muñoz, J. C. (2005). Land surface temperature retrieval from thermal infrared data: An assessment in the context of the Surface Processes and Ecosystem Changes Through Response Analysis (SPECTRA) mission. *Journal of Geophysical Research D: Atmospheres*, *110*(16), 1–10. https://doi.org/10.1029/2004JD005588

Spence, C. O., and Welch, C. D. (1999). *Phosphorus Fertilization for Grain Sorghum Production in the Texas Blacklands*.

Srinivasarao, C., Venkateswarlu, B., Lal, R., Singh, A. K., and Kundu, S. (2013). Sustainable management of soils of dryland ecosystems of india for enhancing agronomic productivity and sequestering carbon. In *Advances in Agronomy* (1st ed., Vol. 121). Copyright © 2013 Elsevier Inc. https://doi.org/10.1016/B978-0-12-407685-3.00005-0

Stafford, J. D., Reinecke, K. J., Kaminski, R. M., and Gerard, P. D. (2006). Multi-stage sampling for large scale natural resources surveys: A case study of rice and waterfowl. *Journal of Environmental Management*, *78*(4), 353–361. https://doi.org/10.1016/j.jenvman.2005.04.029

Struik, P. C., and Kuyper, T. W. (2017). Sustainable intensification in agriculture: the richer shade of green. A review. In *Agronomy for Sustainable Development* (Vol. 37, Issue 5). Springer-Verlag France. https://doi.org/10.1007/s13593-017-0445-7

Swelam, A., Farag, A., Ramasamy, S., and Ghandour, A. (2022). Effect of Climate Variability on Water Footprint of Some Grain Crops under Different Agro-Climatic Regions of Egypt. *Atmosphere*, *13*(8), 1180. https://doi.org/10.3390/atmos13081180

Sylvester, G. (n.d.). *E-AGRICULTURE IN ACTION: BIG DATA FOR AGRICULTURE*.

Tariq, A., Yan, J., Gagnon, A. S., Riaz Khan, M., and Mumtaz, F. (2022). Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. *Geo-Spatial Information Science*. https://doi.org/10.1080/10095020.2022.2100287

Taylor, J. C., Wood, G. A., Earl, R., and Godwin, R. J. (2003). Soil factors and their influence on within-field crop variability, Part II: Spatial analysis and determination of management zones. *Biosystems Engineering*, *84*(4), 441–453. https://doi.org/10.1016/S1537-5110(03)00005-9

Tercan, E., and Dereli, M. A. (2020). Development of a land suitability model for citrus cultivation using GIS and multi-criteria assessment techniques in Antalya province of Turkey. *Ecological Indicators*, *117*. https://doi.org/10.1016/j.ecolind.2020.106549

Thenkabail, P. S., and Wu, Z. (2012). An automated cropland classification algorithm (ACCA) for Tajikistan by combining landsat, MODIS, and secondary data. *Remote Sensing*, *4*(10), 2890–2918. https://doi.org/10.3390/rs4102890

Thiele-Bruhn, S., Bloem, J., de Vries, F. T., Kalbitz, K., and Wagg, C. (2012). Linking soil biodiversity and agricultural soil management. In *Current Opinion in Environmental Sustainability* (Vol. 4, Issue 5, pp. 523–528). https://doi.org/10.1016/j.cosust.2012.06.004

Thomas, A. (2020). *WP/20/95 Improving Crop Yields in Sub-Saharan Africa: What Does the East African Data Say? IMF Working Paper African Department Improving Crop Yields in Sub-Saharan Africa: What Does the East African Data Say?*

Topuz, M., and Deniz, M. (2023). Application of GIS and AHP for land use suitability analysis: case of Demirci district (Turkey). *Humanities and Social Sciences Communications*, *10*(1). https://doi.org/10.1057/s41599-023-01609-x

Tucker, C. J. (1980). A spectral method for determining the percentage of green herbage material in clipped samples. *Remote Sensing of Environment*, *9*(2), 175–181. https://doi.org/10.1016/0034-4257(80)90007-3

Twomlow, S., Mugabe, F. T., Mwale, M., Delve, R., Nanja, D., Carberry, P., and Howden, M. (2008). *Building adaptive capacity to cope with increasing vulnerability due to climatic change in Africa – A new approach*. *33*, 780–787. https://doi.org/10.1016/j.pce.2008.06.048

Tyagi, N. K., Sharma, D. K., and Luthra, S. K. (2000). Determination of evapotranspiration and crop coefficients of rice and sunflower with lysimeter. *Agricultural Water Management*, *45*(1), 41–54. https://doi.org/10.1016/S0378-3774(99)00071-2

United Nations Environment Programme. (2011). *UNEP Annual Report 2011*.

Ustaoglu, E., Sisman, S., and Aydınoglu, A. C. (2021). Determining agricultural suitable land in peri-urban geography using GIS and Multi Criteria Decision Analysis (MCDA) techniques. *Ecological Modelling*, *455*. https://doi.org/10.1016/j.ecolmodel.2021.109610

Vajsová, B., Fasbender, D., Wirnhardt, C., Lemajic, S., and Devos, W. (2020). Assessing spatial limits of Sentinel-2 data on arable crops in the context of checks by monitoring. *Remote Sensing*, *12*(14). https://doi.org/10.3390/rs12142195

Vanhellemont, Q. (2020). Combined land surface emissivity and temperature estimation from Landsat 8 OLI and TIRS. *ISPRS Journal of Photogrammetry and Remote Sensing*, *166*, 390–402. https://doi.org/10.1016/j.isprsjprs.2020.06.007

Verheye, W., Koohafkan, P., and Nachtergaele, F. (n.d.). *THE FAO GUIDELINES FOR LAND EVALUATION*.

Victor, O. K., and Samson, A. O. (2019). An Application of GIS-Based Multi-Criteria Decision Making Approach for Land Evaluation and Suitability Mapping for Rice Cultivation in Oye-Ekiti, Nigeria. *JOURNAL OF AGRICULTURE AND ENVIRONMENTAL SCIENCES*, *8*(1). https://doi.org/10.15640/jaes.v8n1a3

Vinciková, H., Hais, M., Brom, J., Procházka, J., and Pecharová, E. (2010). Landscape Studies Use of remote sensing methods in studying agricultural landscapes – a review. *Journal of Landscape Studies*, *3*(April), 53–63. citeulike-article-id:9838040

Vinciková, H., Procházka, J., and Brom, J. (2011). Timely identification of agricultural crops in the Temelín NPP vicinity using satellite data in the event of radiation contamination. *Journal of Agrobiology*, *27*(2), 73–83. https://doi.org/10.2478/s10146-009-0014-z

Vlassova, L., Perez-Cabello, F., Nieto, H., Martín, P., Riaño, D., and Riva, J. D. La. (2014). Assessment of methods for land surface temperature retrieval from landsat-5 TM images applicable to multiscale tree-grass ecosystem modeling. *Remote Sensing*, *6*(5), 4345–4368. https://doi.org/10.3390/rs6054345

W Clevers, G. P., Biiker, C., C van Leeuwen, H. J., and M Bouman, B. A. (1994). *A Framework for Monitoring Crop Growth by Combining Directional and Spectral Remote Sensing Information* (Vol. 50).

Waldhoff, G., Lussem, U., and Bareth, G. (2017). Multi-Data Approach for remote sensing-based regional crop rotation mapping: A case study for the Rur catchment, Germany. *International Journal of Applied Earth Observation and Geoinformation*, *61*, 55–69. https://doi.org/10.1016/j.jag.2017.04.009

Wambua, J. M., Ngigi, M., and Lutta, M. (2017). Yields of Green Grams and Pigeonpeas under Smallholder Conditions in Machakos County, Kenya. *East African Agricultural and Forestry Journal*, *82*(2–4), 91–117. https://doi.org/10.1080/00128325.2017.1346903

WANG, F., HUANG, J., TANG, Y., and WANG, X. (2007). New Vegetation Index and Its Application in Estimating Leaf Area Index of Rice. *Rice Science*, *14*(3), 195–203. https://doi.org/10.1016/s1672-6308(07)60027-4

Wang, F., Wang, F., Wang, F., Hu, J., Hu, J., Xie, L., Xie, L., Yao, X., and Yao, X. (2020). Rice yield estimation based on an NPP model with a changing harvest index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2953–2959. https://doi.org/10.1109/JSTARS.2020.2993905

Wang, F., Yao, X., Xie, L., Zheng, J., and Xu, T. (2021). Rice yield estimation based on vegetation index and florescence spectral information from UAV hyperspectral remote sensing. *Remote Sensing*, *13*(17). https://doi.org/10.3390/rs13173390

Wan_IEEE_1997. (n.d.).

Waongo, M., Laux, P., and Kunstmann, H. (2015). Adaptation to climate change: The impacts of optimized planting dates on attainable maize yields under rainfed conditions in Burkina Faso. *Agricultural and Forest Meteorology*, *205*, 23–39. https://doi.org/10.1016/j.agrformet.2015.02.006

Wardlow, B. D., and Egbert, S. L. (2010). A comparison of MODIS 250-m EVI and NDVI data for crop mapping: A case study for southwest Kansas. *International Journal of Remote Sensing*, *31*(3), 805–830. https://doi.org/10.1080/01431160902897858

Wardlow, B. D., Egbert, S. L., and Kastens, J. H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, *108*(3), 290–310. https://doi.org/10.1016/j.rse.2006.11.021

Watson, D. J. (1958). The Dependence of Net Assimilation Rate on Leaf-area Index. In *Source: Annals of Botany* (Vol. 22, Issue 85).

Wenner, C. G. (1983). Soil conservation in Kenya. *Ambio*, *12*(6), 305–307. https://doi.org/10.1007/978-1-349-23206-2_6

Wiegand, C. L., and Richardson, A. J. (1990). *(1990) Use of Spectral Vegetation Indices to Infer Leaf Area, Evapotranspiration and Yield: I. Rationale (AJ)*.

Wimalasiri, E. M., Jahanshiri, E., Suhairi, T. A. S. T. M., Mapa, R. B., Karunaratne, A. S., Vidhanarachchi, L. P., Udayangani, H., Nizar, N. M. M., and Azam-Ali, S. N. (2020). The first version of nation-wide open 3D soil database for Sri Lanka. *Data in Brief*, *33*, 106342. https://doi.org/10.1016/j.dib.2020.106342

Wu, W., Yu, Q., You, L., Chen, K., Tang, H., and Liu, J. (2018). Global cropping intensity gaps: Increasing food production without cropland expansion. *Land Use Policy*, *76*, 515–525. https://doi.org/10.1016/j.landusepol.2018.02.032

Xu, Y., Smith, S. E., Grunwald, S., Abd-Elrahman, A., and Wani, S. P. (2017). Evaluating the effect of remote sensing image spatial resolution on soil exchangeable potassium prediction models in smallholder farm settings. *Journal of Environmental Management*, *200*, 423–433. https://doi.org/10.1016/j.jenvman.2017.06.017

Yan, G., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., and Zhang, W. (2019). Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. *Agricultural and Forest Meteorology*, *265*(March 2018), 390–411. https://doi.org/10.1016/j.agrformet.2018.11.033

Yan, L., and Roy, D. P. (2014). Automated crop field extraction from multi-temporal Web Enabled Landsat Data. *Remote Sensing of Environment*, *144*, 42–64. https://doi.org/10.1016/j.rse.2014.01.006

Yin, S., Li, J., Liang, J., Jia, K., Yang, Z., and Wang, Y. (2020). Optimization of the weighted linear combination method for agricultural land suitability evaluation considering current land use and regional differences. *Sustainability (Switzerland)*, *12*(23), 1–26. https://doi.org/10.3390/su122310134

Yu, Y., Wang, J., Liu, G., and Cheng, F. (2019). Forest Leaf Area Index Inversion Based on Landsat OLI Data in the Shangri-La City. *Journal of the Indian Society of Remote Sensing*, *47*(6), 967–976. https://doi.org/10.1007/s12524-019-00950-6

Zhang, H., Chen, J. M., Huang, B., Song, H., and Li, Y. (2014). Reconstructing seasonal variation of landsat vegetation index related to leaf area index by fusing with MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(3), 950–960. https://doi.org/10.1109/JSTARS.2013.2284528

Zhang, S. wen, Shen, C. yang, Chen, X. yang, Ye, H. chun, Huang, Y. fang, and Lai, S. (2013). Spatial Interpolation of Soil Texture Using Compositional Kriging and Regression Kriging with Consideration of the Characteristics of Compositional Data and Environment Variables. *Journal of Integrative Agriculture*, *12*(9), 1673–1683. https://doi.org/10.1016/S2095-3119(13)60395-0

APPENDICES

APPENDIX I : QUESTIONNAIRE

BASELINE QUESTIONNAIRE

INTRODUCTION

Hilda Manzi a Phd candidate of Kenyatta University are conducting a baseline survey. This is academic research work aimed at enhancing the development of data estimation tool for green gram and sorghum under farm field condition. The survey targets to collect information on green gram and sorghum production to improve data estimation under farm field conditions. The information we are collecting is specifically for research, and will not be used for commercial purposes. Further the information will be treated as confidential and will only be used for research. I will take approximately 10 minutes of your time.

Do you agree to participate in this exercise?

Yes

No

If No, what are your reservation?

……

……

……

……

……

Sign……. Identification code……

SECTION 1: LOCATION DETAILS

Date		Start-Time	
Enumerator:		End- Time	
County		Ward	
District / Sub-		Location	
Sub-location		Village	

Grid Reference of Household Using the GPS (UTM)

Northing (Lat)	**Easting (Long)**	**Altitude**	**Time**

SECTION 2.0: CROP PRODUCTION

2.1 Do you grow Sorghum crop?

[1] Yes

[2] No

2.2 Do you grow green gram crop?

[1] Yes

[2] No

2.3 Which season do you grow the crop?

[1] Short rains (OND)

[2] Long rains (MAM)

2.4.1 How frequently have you gown the crop?

[1] Less than 10 years

[2] For 10 years or more

Name of crop	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Season					

APPENDIX II: NACOSTI RESEARCH LICENSE

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Ref No: **755676**

Date of Issue: **26/May/2021**

RESEARCH LICENSE

This is to Certify that Ms.. hilda KALEKYE manzi of Kenyatta University, has been licensed to conduct research in Machakos, Tharaka-Nithi on the topic: CHARACTERIZATION OF SORGHUM AND GREEN GRAM FOR DATA ESTIMATION AND MANAGEMENT USING REMOTE SENSING IN MACHAKOS AND THARAKA-NITHI COUNTIES for the period ending : 26/May/2022.

License No: **NACOSTI/P/21/10911**

755676

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

APPENDIX III: AUTHORIZATION LETTER

KENYATTA UNIVERSITY
GRADUATE SCHOOL

E-mail: dean-graduate@ku.ac.ke
Website: www.ku.ac.ke

P.O. Box 43844, 00100
NAIROBI, KENYA
Tel. 020-8704150

Our Ref: A99/27646/2019

DATE: 27th October, 2020

Director General,
National Commission for Science, Technology
and Innovation
P.O. Box 30623-00100
NAIROBI

Dear Sir/Madam,

RE: RESEARCH AUTHORIZATION FOR MS. HILDA MANZI – REG. NO. A99/27646/19

I write to introduce Ms. Hilda Manzi who is a Postgraduate Student of this University. She is registered for Ph.D. degree programme in the **Department of Agricultural Science & Technology.**

Ms. Manzi intends to conduct research for a Ph.D. thesis Proposal entitled, **"Characterization of Sorghum and Greengram for Data Estimation and Management Using Remote Sensing in Machakos and Tharaka-Nithi Counties, Kenya."**

Any assistance given will be highly appreciated.

Yours faithfully,

KENYATTA UNIVERSITY OFFICE OF DEAN 27 OCT 2020 P.O. BOX 43844-00100 NAIROBI SCHOOL

PROF. ELISHIBA KIMANI
DEAN, GRADUATE SCHOOL

APPENDIX IV: LIST OF PUBLICATIONS

Manzi, K. H., Ngene , S., and Gweyi-Onyango , J. P. (2023b). Use of Satellite-based Leaf Area Index Data for Monitoring Green Gram Crop Growth in Ikombe-Katangi Area, Machakos County, Kenya. *Asian Journal of Advances in Agricultural Research*, *23*(3), 53–63.

https://doi.org/10.9734/ajaar/2023/v23i3468

Manzi K. H., Ngene , S., and Gweyi-Onyango , J. P. (2023a). Use of Satellite Data to Extract pH Values for Crop Planning and Management of Sorghum and Green Gram in Tharaka Nithi and Machakos Counties, Kenya. *Asian Journal of Advances in Agricultural Research*, *23*(3), 25–32.

https://doi.org/10.9734/ajaar/2023/v23i3464

Manzi K.H and Gweyi-Onyango J P (2022). Geoinformation for Land Suitability Modelling for Climate-Smart Farming in Africa. Springer Nature Switzerland AG 2022. A. Kumar et al. (eds.), Agriculture, Livestock Production and Aquaculture, https://doi.org/10.1007/978-3-030-93258-9_9: pp 155-165

Manzi, H., Gweyi-Onyango, J.P. (2021). Agro-ecological Lower Midland Zones IV and V in Kenya Using GIS and Remote Sensing for Climate-Smart Crop Management. In: Oguge, N., Ayal, D., Adeleke, L., da Silva, I. (eds) African Handbook of Climate Change Adaptation. Springer, Cham. https://doi.org/10.1007/978-3-030-45106-6_35