University of Nairobi

Faculty of Engineering

LEVERAGING GEOSPATIAL TECHNOLOGY FOR IMPROVED SMALLHOLDER FARMER CREDIT SCORING

By

Susan Akello Okeyo

(F80/56949/2020)

B.Sc. (Computer Science) (Makerere University), M.Sc. (GIS) (University of Nairobi)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Geospatial and Space Technology of the University of Nairobi

November 2023

DECLARATION

I Susan Akello Okeyo, declare that this thesis is my own work and contains no information or materials previously published or written by other persons without duly acknowledging in the text. This work has not been presented for the award of a degree in any academic institution.

Signature:...... ....Date: 14/November /2023

The research has been submitted for examination with our approval as University supervisors.

Supervisors:

Prof. Galcano C. Mulaku:

Department of Geospatial and Space Technology

Faculty of Engineering

University of Nairobi

Signature...................Date: 16th November 2023

Dr Collins Mwange

Department of Geospatial and Space Technology

Faculty of Engineering

University of Nairobi

Signature:...................Date: 16th November 2023

DEDICATION

To my parents:

Mr. Moses Okeyo Osunga

&

The late Mrs. Imerina Alesa Okeyo

To: The vulnerable small scale farmers of the world.

ACKNOWLEDGEMENTS

Completing PhD is a communal endeavor and I am profoundly grateful to have encountered a great support system during this journey.

First and foremost, I thank God for granting me the strength, guidance and protection throughout my academic pursuit.

My heartfelt gratitude goes to my family for their unending love, encouragement and unwavering support. Thank you Dad, Esther, Jared, Kevin and my special sister Grace for constantly cheering me on.

In a special way, I would like to appreciate my supervisors; Professor Mulaku, thank you for being an integral part of my PhD journey. Your mentorship, unceasing ideas, patience and valuable guidance have tremendously helped me navigate through the challenges of my research and writing of this thesis. Dr Mwange, thank you for your insightful comments, constant encouragements, practical advice and guidance through all the stages of this research. I could not have imagined having better supervisors in my study!

I am grateful to the Department of Geospatial and Space Technology at the University of Nairobi for providing me with the resources and infrastructure to conduct my research.

I wish to immensely appreciate the AI4D (Artificial intelligence for Development) scholarship Fund and Gandhi Smarak Nidhi Fund scholarship for the financial support that made my PhD journey successful.

Lastly, I am indebted to my research assistants and friends who provided me with moral support and helped me maintain good cheer through and through.

ABSTRACT

A small holder farmer is generally understood to be one that farms on a small piece of land, often taken as 2 ha or less, and largely for subsistence; however, this size threshold varies from country to country, depending on the prevailing ecological and demographic conditions; for example in Kenya it is about 0.5 ha. According to the Food and Agriculture Organization (FAO) of the United Nations, there are about 500 million small holder farmers in the world, and in the developing countries, such farmers produce about 80% of the food consumed there; their farming activities are therefore critical to the economies of their countries, and to the global food security. However, these farmers face the challenges of limited access to credit, often due to the fact that many of them farm on unregistered land that cannot be offered as collateral to lending institutions; but even where they are on registered land, the fear of losing such land should they default on loan payments often prevents them from applying for farm credit; and even if they apply, they still get disadvantaged by low credit scores (measure of credit worthiness). The result is that they are often unable to use optimal farm inputs such as fertilizer, good seeds among others. This depresses their yields, and in turn has negative implications for the food security in their communities and in the world, hence making it difficult to realize the UN Sustainable Development Goal No.2 (no hunger). This study aimed at demonstrating how geospatial technology can be used to leverage farm credit scoring for the benefit of small holder farmers. A survey was conducted within the study area to identify the small holder farmers and farms. Further investigation was conducted to establish the extent to which small holder farmers are financially excluded and the results obtained from statistical analysis revealed that indeed the farmers were financially excluded to a large extent. A sample of 101 surveyed farmers was then subjected to credit scoring by machine learning. In the first instance, the traditional financial data approach was used, and the results showed that over 40% of the farmers could not qualify for the credit. When non-financial geospatial data, namely NDVI was introduced into the scoring model, the number of farmers not qualifying for credit reduced significantly to 24%. It is concluded that introduction of the NDVI variable into the traditional scoring model could improve significantly the small holder farmer chances of accessing credit. Possible approaches by which this new model could be fine-tuned have been suggested, and should the model be adopted by industry, the technical and institutional issues that could feature in the implementation are discussed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF CODE SCRIPTS

ABBREVIATIONS

AFC	-	Agricultural Finance Corporation
ALDEP	-	Arable Land Development
ANN	-	Artificial Neural Network
AUC	-	Area Under the Curve
CBK	-	Central Bank of Kenya
CCGS	-	CEDA Credit Guarantee Scheme
CRB	-	Credit Reference Bureau
DFI	-	Development Finance Institution
DRSRS	-	Department of Resource Surveys and Remote Sensing
FAO	-	Food and Agriculture Organization
FICO	-	Fair Isaac Corporation
FITNES	-	Farmer Information Technology Network
GOK	-	Government of Kenya
IV	-	Information Value
LULC	-	Land Use Land Cover
MLFN	-	Multi-Layer Feed-forward Neural network
NACOSTI	-	National Commission for Science, Technology and Innovation
NDB	-	National Development Bank
NDVI	-	Normalized Difference Vegetation Index
PNN	-	Probabilistic Neural Network
RCMRD	-	Regional Centre for Mapping of Resources for Development
ROC	-	Receiver Operating Characteristics curve
SMME	-	Small, Medium and Micro Enterprise
SOK	-	Survey of Kenya
SWOT	-	Strengths, Weaknesses, Opportunities, Threats

USGS - United States Geological Survey

WOE - Weight of Evidence

YFF **-** Young Farmer Fund

1. INTRODUCTION

1.1 Background

Credit scoring is a statistical analysis which is performed by either lenders or credit reference institutions to assess a person's creditworthiness. For lenders, credit scoring is subsequently used to help decide on whether to extend or deny credit to a borrower (Costa *et al*. 2016). For credit reference institutions, credit scoring is described as a means of making a summary of information on a credit application so as to produce a number called a credit score.

Specifically, this study has utilized the following definition of credit scoring: *A credit score is a statistic used to determine a person's ability to pay back a loan*. Schreiner (2003) notes that, a credit score normally ranges between 300-850 with 300 being the absolute lowest and 850 being the best score possible. This scale may vary from one lender to the next, but generally, the higher the credit score number someone has, the easier it is to qualify for a loan. This score is then compared with a predetermined threshold. Theoretically, if the credit applicant's score is greater than the threshold then credit is granted, otherwise credit is denied. In practice the decision to grant or deny a loan using credit scoring is not quite so clear cut. Kaffenberger and Chege (2016) note that there is a grey area surrounding this issue, while questioning why for instance applicants scoring two or three points above the threshold are granted credit, but those scoring just below the threshold can be denied access to credit.

There are three approaches to credit scoring; one is based on a statistical model to predict the probability that a credit applicant will default. The second is based on expert judgment, while the third is based on a combination of the first two. Hence, credit scoring is a credit risk management technique that analyzes the borrower's risk (Gestel and Baesens, 2008).

Farmers use credit from financial institutions to finance production activities on their farms. Financial institutions need to predict the financial sustainability of the enterprise to ensure that the farmer who is borrowing will have the ability to repay the loan. Given the importance of farmers to their national economies and food security, it is important that they receive the support they need. Financing farmers is one of the major challenges since most of the small farmers come from poor households and lack access to credit facilities. This eventually translates to low and poor levels of production and thus compromises their capacity to contribute to food security.

Farmer credit score can be linked to geospatial technology by using geospatial data and tools; especially remote sensing, to assess the credit worthiness of farms. Instead of using collateral like land title deeds, which in most cases is the norm, financial institutions can be able to evaluate the creditworthiness of a farmer using crop data generated using geospatial tools.

The introduction of geospatial crop data into the process of farmer credit scoring has hardly been investigated and this study seeks to fill this knowledge gap. This would in turn open up the possibility of land vulnerable smallholder farmers offering, not their land, but their growing crops as collateral to a potential creditor.

1.2 Problem Statement

A review of literature such as Gestel and Baesens (2008) and Rice (1994), indicates gaps in approaches to credit scoring for farmers, particularly from the aspect of the financial exclusion of some farmers and explicit use of financial data for credit scoring. Financial exclusion has a major impact on small holder farmers in Kenya; it results in limited financial records, including the financial history of the farmers with lending institutions. Credit history is often required by financial institutions to evaluate the risk of a potential borrower before making lending decisions. Most farmers face hindrances when they want to access credit facilities to fund their farming activities. This is mainly because they lack financial history data which is required by financial institutions to compute credit scores for credit risk evaluation.

In retrospect, financial institutions are faced with a challenge of collecting data from farmers in far flung areas that are hard to reach or access is physically limited due to poor road networks. For example, remote areas with poor access infrastructure are hard to reach, also areas with regular insecurity incidents like Turkana and the border between Kenya and Somalia will hinder access by financial institutions.

To overcome these challenges and reduce financial exclusion, this study investigated how non-financial data which is not related to a person's financial activities could be used in computing credit scores for farmers. Generally, there are few studies in the world, among developed countries which have explored the subject under study; specifically, there is no known study in Kenya that has sought to investigate and demonstrate how geospatial technologies can leverage farm credit scoring. Hence to break this cycle, this study seeks to fill these gaps and at the same time answer the question: how can geo-spatial technologies be used to leverage credit scoring for farmers?

1.3 Study Objectives

General objective:

To demonstrate how geospatial technology can be used to leverage farm credit scoring for the benefit of small holder farmers.

Specific objectives:

1. To identify the small holder farms and farmers in the study area.
2. To determine the extent of small holder farmer financial exclusion.
3. To develop a new farmer credit scoring system that includes geospatial technology.

1.4: Study Conceptual Model

The study was conceptualized as leading to the generation, using a machine learning approach, of an improved credit score for a financially excluded small holder farmer; this improved score is the dependent variable. Contributing to this generation are a number of independent variables, principally the traditional score model plus geospatial data in the form of satellite imagery, farm GNSS positions and NDVI. This conceptual model is illustrated in Figure 1.

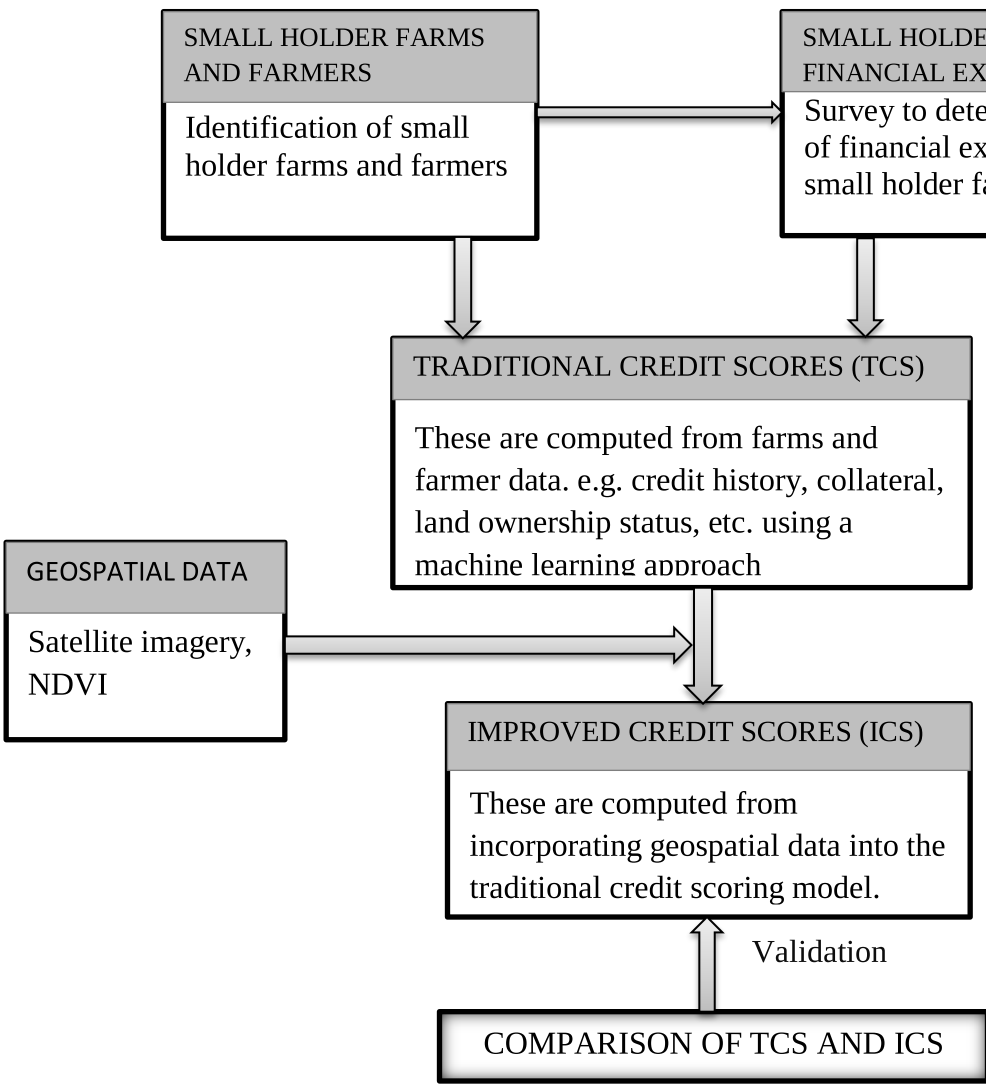

Figure 1: The study conceptual model

1.5: Organization of the Thesis

This work is organized into six chapters informed by the research objectives.

Chapter One: This is the introductory part of the study which contains the background of the study, problem statement, research objectives, study conceptual framework and the organization of the thesis chapters.

Chapter Two: Reviews the literature focusing on the concept and development of credit scoring, current status of credit scoring both locally and globally and what previous studies have been done on this subject.

Chapter Three: Describes in detail the methodology used in this study. It describes the data collection process, data analysis and the variables that were selected to be used in the final credit scoring model; outlining the step-by-step processes which were undertaken in implementing credit scoring using WOE and Generalized linear model in the R programming language.

Chapter Four: Assesses the results obtained from both the traditional scoring model and the new scoring model and provides a comparison of the two.

Chapter Five: Proposes an implementation framework that would fine tune the model and also address the implementation issues such as policy issues, institutional issues and technical issues.

Chapter Six: Gives the conclusions, recommendations and a summary of the key contribution of this study; highlighting areas with possible opportunities for future research.

2. LITERATURE REVIEW

2.1 Concept and development of credit scoring

The emergence of credit scoring can be traced back to 1941 when David Durand established a credit scoring system. Durand identified different variables that helped lenders distinguish between good and bad loans (Gutiérrez-Nieto *et al.*, 2016). In 1946, E. F. Wonderlic developed a credit score guide that helped define and narrow the variables of good and bad loans (Weston, 2012). The credit score guide helped to indicate the degree of risk associated with a customer. In the 1950's credit scorecards were becoming a popular instrument used in credit worthiness assessment.

The Fair Isaac Corporation (FICO) was founded in 1956 in the US and introduced its first credit scoring system in 1958 (myFICO). These scorecards were models that helped to determine if a customer will default on their loan given their current financial position. The late 1960s to early 1970s brought about technology that allowed for credit scoring models to be developed further and automated (Thomas, *et al*, 2004).

Credit can be defined as control over money, materials, goods or services in the present in exchange for a promise to repay at some future date (Lawal *et al.*, 2009). This implies that, lenders forgo the use of money or its equivalent in the current time by making loans available or extending the credit to the borrower who promises to repay on terms specified in the loans agreement or debt instruments (Barry and Robinson, 2001). It is an advance of money or its equivalent given by a lender to a borrower for repayment at maturity, which may range from a few days to several years (Llanto, 2005).

A credit risk score is a number indicating the probability of a person paying what they owe; it is produced by evaluating information from one's credit reporting agency. As a system, credit scoring was used for the first time in the United States by retailers and mail-order companies in the 1950's to manage and diversify borrowers default risks (Farrin and Miranda, 2015). Since then, the use of credit scoring has evolved in recent times and is used in banking and finance circles.

On the one level, credit scoring uses the borrower's historical data and credit characteristics to detach the effect of several characteristics of applicants on defaults. On the other level, credit scoring analyzes electronically the borrower's credit history and other characteristics regarding repayment ability that are, in general, provided by borrowers. Based on previous experience with borrowers of similar loan profiles, credit models could therefore predict the default risk of any loan granted. Subsequently, a successful credit model should give high scores to borrowers whose loans would perform well and low scores to borrowers whose loans would not perform well.

There are two broad means of evaluating credit worthiness: appraisal of repayment capacity and asset backed lending. The first approach focuses on investigating the integrity, moral character, management ability and debt repaying capacity of a potential borrower either through human experts or statistical models, while the latter focuses on the quality and quantity of assets that can be mortgaged or pledged as collateral and quickly liquidated in the event of default (Peck *et al*. 2013).

The decision-making process for credit scoring can be either subjective or objective (statistical) (Schreiner 2003). Subjective scoring relies on the input of an expert, the loan officer, and the organization to produce a qualitative judgment. Statistical scoring, on the other hand, relies on quantified characteristics of the prospect's portfolio history recorded in a database. It uses a set of rules and statistical techniques to forecast risk as a probability.

The relationship between risk and client characteristics is expressed as a set of rules in a mathematical formula that forecasts risk as a probability (Rice, 1994).

$$Z_i = ln\left(\frac{Pi}{1-Pi}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$$

Where, p_i is the probability that an individual access credit given X_i. X_i Represents the i^th explanatory variables; α and β_i are parameters to be estimated. Central to the use of the logistic regression is the logit transformation of p given by Z.

Credit scoring not only tells if the client is risky or not; it also provides a measure of the predicted risk. Credit scoring analyzes the characteristics and performance of past loans to predict the performance of future loans.

2.2 Other approaches to credit scoring

2.2.1 Neural networks

A neural network is a collection of neurons that take input and, in conjunction with information from other nodes, develop output without programmed rules (Piramuthu, 1999). The network makes decisions by assigning each connected node to a number known as a "weight." The network gives more weight to data that supports correct guesses and less weight to data that leads to mistakes. A feature known as back propagation trains the network to identify correct responses and ignore incorrect responses. The functions of a neural network are to score inputs, calculate loss and update the model, which begins the process over again (Jensen, 1992). Neural networks excel at classification tasks, which require labelled datasets for supervised learning.

While they excel at identifying differences, neural networks also work well for clustering or detecting similarities. This capability is useful for identifying anomalies, or things that don't correspond with group characteristics. For example, clustering is used to identify unusual behaviour such as fraud by identifying data that doesn't correspond with the most common actions. Artificial neural networks (ANN) are used to forecast the occurrence and extent of spatial events (Lacher et al, 1995). ANN is applied during training and testing. Training data are used to derive the relationships between the dependent variable and the controlling parameters.

How ANN can be applied to Credit scoring

Credit classification techniques are usually estimated through three properties, namely: accuracy, interpretability, and computational efficiency (Mester, 1997). Accuracy is the essential requirement which represents that the maximum possible number of correct decisions can be generated. And a minor improvement of accuracy means a significant saving for a financial institution. The interpretability is quite important to not only decision makers but also credit applicants, since it represents the ability to generate an understandable evaluation mechanism to the applicants, which includes the choice of the most essential input attributes of the analysis model in the meantime.

The computational efficiency represents the speed of classification. It is helpful for the assessors to make the decision as to whether credit should be granted or not as quickly as possible according to the classification result (Henley, 1995). Therefore, the credit classification model which owns the above-mentioned properties can be considered as an appropriate tool in the business and finance fields, especially under the conditions with high uncertainty.

2.2.2 Decision trees

A decision tree is a classification procedure that recursively partitions a data set into smaller subdivisions on the basis of a set of tests defined at each branch (or node) in the tree. It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure (Alpaydin, 2010). In order to build a tree, a CART (Classification and Regression Tree) algorithm is used**.** A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into sub trees (Quinlan, 1986).

There are three steps involved in the building of a decision tree (Vens *et al*, 2008). Splitting is the process of partitioning the data set into subsets; splits are formed on a particular variable. Only input variables related to the target variable are used to split parent nodes into purer child nodes of the target variable. In pruning; classification trees may fit the training data well, but may do a poor job of classifying new values. In tree selection, the process entails finding the smallest tree that fits the data. Usually this is the tree that yields the lowest cross-validated error.

Applications in the geospatial field include analysis of groundwater productivity-potential and examining the method of decision tree for spatial data classification (Schleiter *et al*, 1999). In credit scoring, a decision Tree based model can handle credit granting decision support system using an integration of Decision Tree and Artificial Neural Networks with a hybrid of Decision Tree algorithm and Multilayer Feed-forward Neural Network with back propagation learning algorithm to build up the proposed model.

A classification model based on a decision tree by learning historical data can be used to improve credit scoring (Fielding, 1999). Clustering algorithm and genetic algorithm can be combined to improve the accuracy of this credit scoring model. In this case, the clustering algorithm would aim at removing noise data, while the genetic algorithm would be used to reduce the redundancy of attribute data. The computational results on the two real world benchmark data sets showed that the presented hybrid model was efficient (Blookeel *et al*, 1998).

2.3 Credit scoring and farm credit

In farming, credit intervention is considered to be an effective tool to eliminate poverty. Credit is a major part of financial capital that can assist farmers to benefit from financial resources which are beyond their own capabilities, and therefore making it easier to take advantage of possible opportunities that would be profitable to their business (Zellar and Sharma, 1998).

Farmers can use credit from commercial credit providers to finance production activities. For this to be achieved, the credit providers must be able to predict the financial sustainability of the farming business in order to be sure that the loan will be repaid by the farmer (Hananu *et al*, 2015). In farming, the ability to raise capital is vital and access to credit is a major boost to farmers in terms of reaping maximum returns. However, most farmers are unable to access this credit despite its importance.

When creating a credit scoring service, the scoring algorithm, methodology, and processes used depend on an organization's objectives. A number of questions need to be addressed here. For instance; if the scoring solution is for a new or for an existing loan product, if the scoring solution is for a new or an existing customer, if there are previous loan performance data available, if the data is reliable, updated, and accessible, if there are any other external credit data sources, and if there are any non-financial external data sources.

In most cases, small holder farmers are disadvantaged because their lands are not titled; this means that no financial body can deal with them. Titled land translates into more credit to farmers which enables them to invest in the land and hence higher income for farmers.

Feder's conceptual model (see Figure 2) illustrates how titled land provides more security to the farmer, since it can be used as collateral for getting a loan; this will amount to more demand for investment, leading to more investment in the farm. On the other hand, titled land provides security to the lender, in the sense that it can be used to pay back loans in cases of defaults; this will encourage more supply of cheap long term credit, leading to more investment. More investment means more demand for variable inputs for farmers and more variable input use. Hence achieving higher output per acre, this will lead to higher income for the farmers and also increased value for the land.

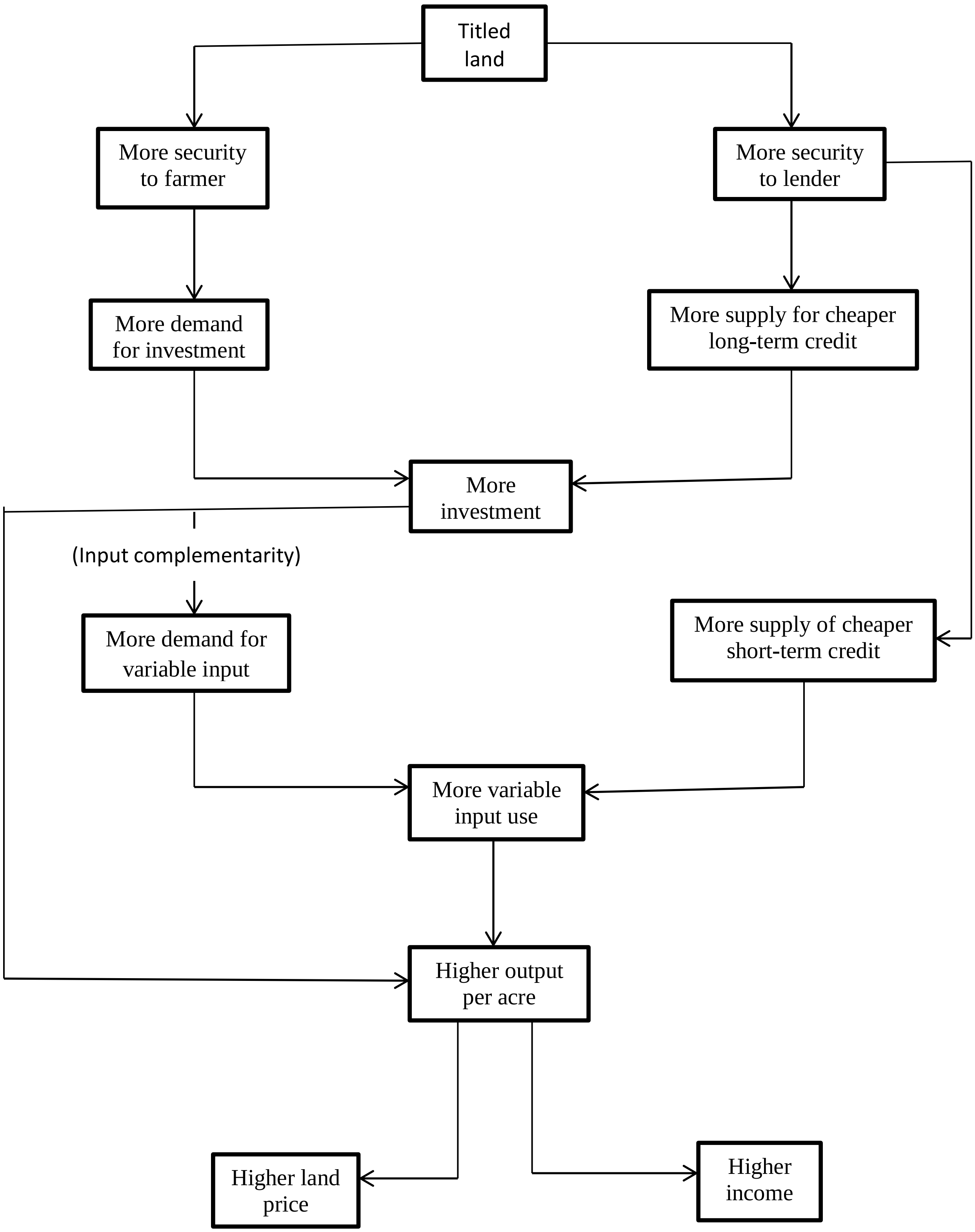

Figure 2: Land ownership security and farm productivity conceptual framework [Feder, (1987)]

Access to credit facilities by small and poor rural farmers has a two-fold advantage; first, access to credit has the potential of making the difference between poverty and economically secured life; secondly, access to credit can enhance agricultural productivity of farmers (World Bank, 2016). Conversely, limited availability of credit can undermine income activities for farmers due to lack of capital for investment and can prevent farmers from adopting improved farming practices. Lack of credit facilities has been regarded as the major constraint which farmers face when they try to increase their economic activities and or living conditions. (Udry, 2000).

Credit accessibility is important for the improvement of quality and quantity of farm products, so that it can increase a farmer's income and avoid rural to urban migration (Siles *et al* 1994). Credit accessibility is the ease or difficulty of acquiring credit by borrowers for purposes such as to enhance business performance. Lenders place significant weight on the borrower's financial information and personal characteristics (honesty, integrity, and production-management ability) when making decisions regarding approval, levels of credit, and need for servicing action (Gustafson *et al*. 1991).

2.4 Credit scoring: The place of geospatial technology

The application of Geospatial technology is relatively new to credit scoring. The financial industry is on a steep learning curve on how to best utilize geospatial technology. There are a number of studies examining the agricultural lending decision making process. Mc Evoy (2013) provides strong evidence that lenders consider both financial and non-financial variables when evaluating the credit-worthiness of farm borrowers. Furthermore, while there have been many studies, such as Liverman *et al* (1998), and McEvoy (2013), the majority of them do not explicitly state how lenders could use geospatial technology when assessing farm borrowers.

Geospatial technology, especially remote sensing, can be applied in farmer credit assessment to objectively assess the health of crops, through NDVI measurements. Higher resolution satellite imagery (such as Landsat TM and Sentinel-2) can be especially suited for this purpose as they better enable the resolution of the small farms associated with smallholder farmers. NDVI measurements would enable not just crop health assessment, but also estimation of the expected yield at harvest. Such yield estimates, together with traditional financial data, could be used to compute a credit score that is friendlier to smallholder farmers. Liverman *et al* (1998) observes that geospatial technology compliments and enhances the traditional subjective assessment of credit scoring among farmers; since it measures credit risk of the applicants quickly and more accurately compared to the scoring methods discussed earlier.

2.5 Credit Scoring in Kenya

There are three main actors involved in credit scoring for farmers in Kenya; insurance companies, commercial banks, and the Agricultural Finance Corporation (AFC). Donor-sponsored organizations such as Syngenta Foundation also complement credit scoring mechanisms both directly and indirectly (Ngare *et al.*, 2015).

2.5.1 Agricultural Finance Corporation (AFC)

AFC is a wholly owned Government Development Finance Institution (DFI), established through an Act of Parliament (Agricultural Finance Corporation Act, Cap 323 of the Laws of Kenya) to provide credit facilities geared towards developing agriculture. AFC employs three techniques in assessment and evaluation of the viability of credit proposals; CAMPARI, 5Cs (Character, Capacity, Capital, Condition and Collateral) and SWOT. These techniques provide an objective approach by which sufficient information about all the relevant aspects of a business can be gathered and analyzed so that a decision can be made on the best available evidence. CAMPARI and 5Cs focus on the factors that are to be analyzed while SWOT addresses the wider issues which affect the performance of a business (AFC Kenya, 2018). AFC has institutionalized the techniques through a system known as FITNES, an acronym for Farmer Information Technology Network Enterprise System, an innovative cloud computing model that integrates real time credit delivery with agronomic dynamics on the field. The effect of the system has been to reduce the default rate from above 50% to below 30% by having less subjectivity in decision making.

Credit scoring in AFC is systematically exercised via two credit products; the cash crop loan and the horticultural and floricultural loan. The Cash Crop Loan targets cash crop production and improvement. It covers cash crops such as tea, coffee, sugarcane, and bananas. The loan finances crop establishment, crop maintenance, processing equipment, and operating costs. Repayment of this loan ranges between two to five years by installments and is designed for individuals and groups. For farmers to be eligible, they must have tangible security for the loan, appropriate and approved crop varieties, and availability of processing facilities within reasonable distances.

To qualify for credit from AFC, farmers are required to provide viable proposals, with complete plan on how to implement their projects. If approved by AFC upon assessment, they will be given the funds which will cater for set-up and operation capital.

2.5.2 Commercial Banks

Kilimo Biashara loan is a product offered by Equity Bank for agricultural credit products. Its purpose is to finance purchase of farm inputs such as fertilizers, certified seeds, machinery hiring chemicals, labor and harvesting costs, for its target beneficiaries, who are the small scale commercial food crop farmers. For one to be eligible for this credit facility, an applicant needs to fulfill the following conditions: One should be an active account holder with Equity Bank, loan applications should be submitted within a reasonable time (preferably one month) prior the setting in of planting season, demonstrate strongly ability to repay, be willing to attend group meetings weekly, as per the agriculture group lending policy and be in commercial farming with farming experience of at least one successful season.

In addition to that, the farmer should be able to demonstrate clearly the existence of other sources of income that could be used to pay the in case of crop failure, loss of harvest and/or poor marketing due to adverse weather and any other factors; Provide evidence of ownership of the land/farm to be used for the production or a valid lease agreement covering at least two future seasons; the same must be signed and witnessed by a lawyer; Identify inputs suppliers, negotiate the inputs prices and obtain quotations/pro-forma invoices for the inputs to be financed; Be able to demonstrate understanding of the market for the commodity being financed. (Salami *et al*, 2010)

2.5.3 Jamii Bora Agribusiness Loan

This credit product by Jamii Bora bank is specially designed to cater for Agri- entrepreneurs and farmers involved in agricultural related production and agribusiness activities. These include but are not limited to horticulture, floriculture, and livestock farming such as poultry, piggery, beef farming, apiculture and aquaculture. It extends a credit amount in the range Kshs. 50,000 to Kshs. 3,000,000 whose loan requirements are: An account holder with Jamii Bora Bank; Owner of an Agri-enterprise for at least one year; Small holder farmer should be a land or lease owner with one successful crop cycle; Credit/Insurance required on Loan. The benefits of these credit facilities are mentioned as low interest rates, flexible collateral requirements and a loan up to 24 Months. (FSD Kenya Report, 2018)

2.6 Credit scoring elsewhere in Africa

2.6.1 Botswana

There are two guarantee schemes in Botswana with a mandate to provide loans for farmers; the CEDA credit Guarantee Scheme (CCGS) and the Young Farmers Fund (YFF) (National Development Bank, 2006). The projects covered under both of these schemes include livestock, crop production, horticulture, agricultural machinery and equipment, agricultural inputs such as fertilizers, chemicals, seeds, feed, labor, contract ploughing, land development, among others.

CEDA Credit Guarantee Scheme (CCGS) is the program launched by the Government of Botswana in an attempt to address the lack of access to credit and the inability of the small and medium scale enterprises (SMMEs) to fulfill the security requirements of financial institutions. The CCGS was created as an effort to motivate and encourage both the commercial banks and other private sector financial institutions to embark on the development of the SMMEs.

Through this program, in 2005, the Government of Botswana decided to create the Young Farmers Fund (YFF) as a fund under the umbrella of CEDA to provide agricultural loans to young people aged between 18 and 35 years. Young farmers can obtain access to credit and entrepreneurial training, so that they can engage in sustainable agricultural activities after having been better equipped with the required skills for running a farming business (CEDA, 2006).

In an attempt to improve access to credit for smallholder farmers in Botswana, the government reformed its NDB to make it a more viable financial institution. For example, in the early 2000s, the Government of Botswana suspended agricultural subsidy schemes. The government also partnered with farmers in managing and financing such schemes by means of the Arable Lands

Development Program (ALDEP) and the Livestock Management and Infrastructure Development Project (LMIDP) (Ministry of Agriculture, 2006).

The LMID program was established with these primary objectives: to promote food security by means of improved productivity of cattle and small stock, improve livestock management, and range resource utilization and conservation. The program was also aimed to eliminate destitution by providing resources to the poor, including infrastructure for hygienic and safe processing of poultry products. The loan application procedure for cattle owners requires that they should possess proof of ownership from the local extension agents and headman and a registered cattle brand. Applicants have to state the current numbers of livestock owned, and verified by the local extension agent. Ten years is regarded as the lifespan of each project and during that period farmers should submit returns on the number of their cattle to the extension agent every year. Applications should be submitted to district offices across the country, for the district officer to verify that the necessary information has been provided. If the application complies with the selection criteria, it is submitted to headquarters for approval. A desk office is established within the Department of Animal Production to receive and process applications and a committee is established to receive and process applications. Applicants are informed in writing whether their applications have been approved or rejected and a copy of this letter is sent to the district office (Ministry of Agriculture, 2006).

Despite the fact that the documentation sources for the case of Botswana could not provide any performance indicators of the NDB, it was to some extent, useful in showing the role governments need to play if improving access to credit for smallholder farmers is to materialize as a strategy to boost the agricultural sector of the country. The credit schemes under CEDA achieved considerable gains in terms of improving access to credit for smallholder farmers.

These improvements can particularly be attributed to their special focus on the development of viable and sustainable citizen-owned business enterprises.

2.6.2 Zimbabwe

Since independence, the Government of Zimbabwe has always played an active role in extending agricultural credit to smallholder farmers for crop and livestock production. Rukuni, *et al*, (2006) note that in 1980 and 1981, the government declared that loans would be granted to smallholder farmers in an attempt to redirect institutionalized agricultural credit. The smallholder farm credit scheme was thus established through two statutory financial instruments. This led to the establishment of the Agricultural Finance Corporation (AFC) which granted credit to smallholder farmers (Zumbika, 2000). Access to credit for smallholder farmers improved while long-term loans were granted to large-scale commercial farmers for infrastructural development projects (Rukuni, 2006).

Immediately after the independence of Zimbabwe in 1980, the AFC was given a new mandate to shift its focus from lending to white commercial farmers to smallholder farmers. From 1980, all loans at the AFC were extended to smallholder farmers. In order for the AFC to successfully perform this role, the government guaranteed all loans to smallholder farmers. The AFC indeed continued to perform this role for almost two further decades from 1980 until 1999 (Zumbika, 2000).

Credit facilities were provided in cash for the purchase of agricultural inputs, farm equipment, machinery and livestock, as well as for the purchase of a farm and working capital. From 1982 up to 1997, the government ensured that the AFC was well-funded to meet the primary objectives of extending and improving access to credit for smallholder farmers. The strategy by the Government of Zimbabwe to improve access to credit for smallholder farmers by promoting cash crops such as tobacco, led to the diversification from flue-cured tobacco, only grown by white commercial farmers, to other types of tobacco (air-cured and oriental) which mainly smallholder farmers began to grow.

Group Lending Methodology

In order to render the smallholder credit system more sustainable, a group lending methodology was introduced as a new requirement for the smallholder farmers (Chigumira and Masiyandima, 2003). The group would be structured to elect a group committee composed of a chairman, secretary, and treasurer. The responsibility of this committee would be to mobilize loan repayments by the group members in order to repay the AFC, while the treasurer would be responsible to liaise with the AFC official to negotiate a group loan.

Conditions of access to agricultural credit at Agribank

According to Sacerdoti (2005), the following are the conditions for farmers, including smallholder farmers, to access credit from the Agribank of Zimbabwe: proof of legal age to borrow and identification of particulars; proof of land ownership: communal land holders and rural resettlements; program of action: the farmer indicates the size of land, what crops he/she wants to grow and other relevant information on the application form; and a credit track record: the farmer has to explain how he/she has been performing in terms of credit repayment.

The case of Zimbabwe demonstrates that direct government intervention in rural financial markets succeeded in its specific mandate of improving access to credit for smallholder farmers. It can be concluded that when appropriate strategies are implemented and the government takes the lead in the process, improved access to credit for farmers can be achieved. Thus, the success of the AFC in improving access to credit for smallholder farmers may be attributed to the role played by the government of directly intervening in the rural credit markets through establishing the credit institution, mobilizing funds and financing it to enable it to fulfill its mandate. In addition, the government also established and invested in other appropriate institutions which complemented the role of AFC and ensured that practical farm problems were addressed.

2.6.3 Credit scoring in Ghana

The history of credit bureau agencies in Ghana can be traced back to the early 2000s when the need for structured credit information became apparent. With a growing economy and expanding financial sector, the limitations of informal credit assessments began to surface. Most banks in Ghana use lending as a major source of income. Banks allocate sizeable portions of their assets into varied types of credit. Lending is done primarily with the intent of generating enough income to contain expenditure and enhance shareholder value. Current competition within the Banking Industry is compelling Banks to either be cautious with their lending activities or to be flexible with their lending criteria so as not to lose their client base (Quainoo, 2011). In competitive periods in Banking, Banks to a large extent are less cautious and more fluid in advancing credit (Mensah, 2004). Establishing formal credit bureau agencies has marked a significant step towards a more transparent and efficient credit market in Ghana.

Ghana has two licensed credit bureau agencies that offer comprehensive services to financial institutions and businesses. They support various financial products and services catering to Ghana’s population’s diverse needs.

The functions of those credit reference bureaus are:

- Collecting and Maintaining Credit Information
- Analyzing Creditworthiness
- Facilitating Credit Access for Individuals and Businesses

- supporting Banks and Financial Institutions in Risk Assessment

2.7 Credit scoring elsewhere globally

2.7.1 India

The practice of agricultural credit scoring in India is arguably the oldest in the globe. It was founded at a time when governments, the world over, were the ultimate overseer of state affairs, and in India, little has changed since then. The second milestone of agricultural credit scoring happened in the mid-20th century when the State Bank of India was created. The state bank extended banking services to the rural areas of the country.

The foundation of agricultural credit scoring is founded on the Government's attempts to eradicate the problem of rural indebtedness. The initial step taken by the Government of India towards addressing the problem was the establishment of co-operative credit societies (Saima and Hussain, 2011). Subsequently, the Co-operative Credit Societies Act (1904) was passed with the aim of providing cheap and cost-effective financial services to Indian farmers. Under the concept of priority sector lending, commercial banks were mandated to advance a certain proportion of their funds to the agriculture sector.

There were three rife factors which necessitated the need for agriculture credit: increased gap between income and expenditure; unevenness of investment in fixed capital formation; and surges in capital needs (Chavan and Ramakumar, 2007). The Reserve Bank of India moved to improve the availability of farm loans from commercial banks to neglected areas in 1972 by introducing the requirement that banks allocate a proportion of aggregate bank advances to priority sectors. The banks were advised to increase their share to 33.5 per cent of adjusted net bank credit (ANBC) or the credit equivalent amount of off-balance sheet exposure (OBE), whichever is higher, by March 1979 (Satyasai, 2012).

For almost 70 years, co-operatives were the main instruments for extending agricultural credit in India. However, this changed in 1969 with a shortfall in agricultural output in India; the 'All-India Rural Credit Review Committee' then recommended the adoption of a 'multi-agency approach' towards agriculture and rural credit. The Narasimham Committee on rural credit (1975) recommended the establishment of Regional Rural Banks, as it was of the view that neither commercial banks nor co-operative institutions were able to meet agricultural credit needs.

Another milestone on agricultural credit scoring in India came about in the 1990s with the commissioning of the three other initiatives; the Kisan Credit Card Scheme, Self Help Group-Bank Linkage Program and Special Agricultural Credit. The three were put in place to increase the flow of credit to the agricultural sector. With a view to ensuring that the flow of credit to agriculture increased, the Reserve Bank of India (RBI) advised banks in 1994-95 to prepare an action plan for disbursement of credit to agriculture. Accordingly, each bank prepares a Special Agricultural Credit Plan (SACP), segregated into quarterly targets, which is monitored by the RBI. Earlier, the SCAP mechanism was applicable only to the public sector banks but it was extended to private sector banks in 2005-06.

Agricultural credit plan in India

Agriculture credit can be classified into long-term and short-term credit. Short term credit is typically for six months but covers credit up to one year. Long-term credit includes the medium term. Direct lending to farmers takes the form of either short-term or long-term credit (Pal, Sapre, 2010). Short-term agricultural credit was crafted to enable cultivators procure inputs such as fertilizer and seeds needed a planting season. Short-term credit is also meant to cover the cost of hired labor as well as a part of the consumption needs of poorer farmers. In comparison, long-term credit is extended for investment in fixed assets, for instance, irrigation pumps, tractors, agricultural machinery, plantations and those related to dairying, fishing and poultry.

ICICI Bank Satellite imaging

Since March 2020, ICICI bank of India commissioned a satellite imaging system to measure an array of parameters related to the land, irrigation and crop patterns, in combination with demographic and financial markers, to help it in making lending decisions (ICIC, 2020). The bank partnered with agri-fintech companies specializing in harnessing space technology and weather information for commercial usage. The analysis is put together using algorithms to scour images available from satellites around the planet.

ICICI has worked on further scoring models to create indices at district level, village level as well as for individual land to provide an estimate of the past and future agriculture income, the timing of harvest and sources of income, to deliver detailed inputs to credit assessments.

Harvesting Enterprise

Harvesting is a United States Based company with a presence in India. They specialize in agricultural credit scoring for small holder farmers (Virendra, 2018). The approach used by harvesting enterprise uses remote sensing data from NASA and EU Space Agency satellites and combines it with AI-backed algorithms to derive a credit score, thus helping financial institutions to arrive at a data-driven decision.

Harvesting has been in operation in India since 2018, they built an Agricultural Intelligence Engine (AIE) as its core technology, which applies big data analytics and uses remote sensing satellite images and Machine Learning (ML) to design farmers' credit score and provide data points to financial institutions. It utilizes AI remote sensing tools to bridge the data gap for farm credit in emerging markets.

It has developed global crop identification analytics and metrics to understand productivity and growth (such as vegetation cover, water stress, and diseases). Combined with geo-spatial factors (e.g. climate, topography), historical financial data, value-chain data (e.g. buyer contracts), and transaction data. Harvesting, with its predictive ML algorithms, created a product for financial institutions to provide them specific farm land data by putting in farmer's survey ID. Then it gives a snapshot of the land, its yield capacity and types of crops that can be cultivated, thus making it more efficient for banks to disburse loans.

In conclusion, it is apparent that directed lending by commercial banks envisaged under the priority sector lending mandate has been a critical factor contributing to the expansion of agricultural credit in the country. It must be acknowledged, however, that the inclusion of indirect financing within the scope of lending to agriculture would have released the pressure on banks for direct lending to farmers.

2.7.2 USA

Agricultural credit scoring in the United States of America was started to provide support for farming because of its significance to the well-being of the U.S. economy. The Farm Credit System (FCS) was established to provide a permanent, reliable source of credit to U.S. agriculture in 1916 (Monke, 2015). The FCS has a statutory mandate to serve agriculture, and certain agribusinesses and rural homeowners. Borrowers must meet certain eligibility requirements in addition to general creditworthiness.

Agriculture credit scoring is carried out by three main organizations: The USDA Farm Service Agency (FSA); The Farm Credit System (FCS) and Farmer Mac c. Other sources of credit for the agricultural sector include commercial banks, life insurance companies, and individuals, merchants, and dealers.

The US Farm Credit System has collaboration with some institutions, and this system is controlled by the Congressional Agricultural Committee. Farm Credit System was established by Congress in 1916 to provide a dependable and affordable source of credit to rural areas at a time when commercial lenders avoided giving farm loans. The Farm Credit Administration sets minimum regulatory capital requirements for banks and associations.

Generally there are three major US credit bureaus: Equifax, Experian and Transunion. The credit bureaus gather information on one's credit use and provide it to lenders and other businesses. Information about credit practices, such as paying bills on time, credit limits, number of accounts, amount of debt, types of credit accounts, etc. is used to determine your credit score.

There are two scores that combine information from the three agencies – FICO and Vantage Score; although many lenders use the credit score generated by the Fair Isaac Corporation called FICO which ranges from 300 to 850 (www.annualcreditreport.com).

FICO score components and general percentages of importance are:

•**Payment History (35%)**– Regular payments made on time

•**Amounts Owed (30%)**–Total amounts owed and ratio or balances to available credit

•**Length of Credit History (15%)** – Length of time credit has been used, length of time accounts have been open

•**New Credit (10%)** – Number and type of new accounts, account inquiries

•**Types of Credit in Use (10%)** – Variety of types of credit including credit cards, installment loans, mortgage, etc.

2.8 Previous studies on Credit Scoring

2.8.1 Case study of Thailand

In Thailand, an analysis of credit scoring for agricultural loans was conducted in 2005 (Limsombunchai and Minsoo, 2005). The aim of conducting that study was to find out how to estimate a credit scoring model that would be used in making lending decisions for the agricultural loans. Probabilistic Neural Network (PNN) which is a special class of artificial neural networks was used to estimate the credit scoring model together with the logit model and an Artificial Neural Network (ANN) called Multi-Layer Feed-Forward Neural Network (MLFN)". The study compared the predictive power among the three different estimation methods.

The scoring model was:

Lending decision = f (Borrower characteristics, Credit risk proxies, Relationship indicators, Dummy variables)

Where, Lending decision =0 if loan is defaulted on (bad loan); 1 of loan is paid (good loan)

Borrower characteristics: Assets, age, education; where 0 if the qualification if the borrower is primary school or lower; 1 otherwise;

Credit risk proxies include: Collateral, Return on assets, Leverage ratio, Capital turnover ratio;

Relationship indicators: Borrowing from others (1 if the borrower has debt with other financial institutions and 0 if the borrower has a loan from only BAAC), Duration (duration of bank-borrower relationship);

Dummy variables: Loan size, Province, Loan type, Farm type, and Lending year.

Dummy variables are useful when describing the systematic effects directly relating to the borrowers and the type of contracts. For instance, borrowers who produce cash crops tend to require less amount of credit compared to other farm types. This makes their contract term short, making these borrowers obtain a higher probability of being granted a loan than the others; the reason being that short-term loans are less risky than medium-term or long-term loans and the lending risk is relatively low.

MLFN model: The ANN model can be represented as a massive parallel interconnection of many simple computational units (or neuron or node) interacting across weighted connections (Venugopal and Baets, 1994). Each neuron or node is made up of a set of input connections that receive signals from other neurons, transfer function and a set of weights for input connection. The output for node j, U_j, is the result of applying a transfer function F_j to the summation of all signals from each connection (X_i) times the value of the connection weight between node j and connection I (W_{ij})

$$U_j = F_j \left(\sum W_{ij} X_i\right)$$

Where, U_j is the output for node j and F_j is the transfer function which contains different functional forms: linear threshold functions, Gaussian function, linear functions, sigmoid functions or step functions.

The (MLFN) computational units are categorized into three main layers: the input layer (first layer), the output layer (last layer) and the layer(s) in between (hidden layer). Since output of one layer is an input to the next layer, the output of the network (Z) can be mathematically presented as follows:

$$Z = F\left[\sum_{j=1}^{J} W_j^{(2)} . F_j \left[\sum_{i=1}^{i} W_{ij}^{(1)} X_i\right]\right]$$

Where, Z represents the output of the network, F is the transfer function in the output node, $W_{ij}^{(1)}$ and $W_{j}^{(2)}$are connection weights from input layer (node i) to hidden output layer j and from hidden layer (node j) to output layer, respectively.

PNN model: The PNN model proposed by Specht (1990) is a classification network with a general structure consisting of four layers: a pattern layer, an input layer, an output layer and a summation layer. This model is based on the Bayesian classifier statistical principle. According to the theorem, X will be classified into class A, if the inequality in holds for the following equation:

$h_A c_A f_A (X) > h_B c_B f_B (X)$

Where, X is the input vector to be classified, h_A and h_B are prior probabilities for class A and B, c_A and c_B are costs of misclassification for class A and B, $f_A(X)$ and $f_B(X)$ are probabilities of X given the density function of class A and B, respectively (Albanis and Batchelor,1999).

PNN model working principle begins from the first layer (input layer), which distributes the inputs to the pattern layer. Then the pattern layer memorizes each training sample and estimates the contribution of a particular pattern to the probability density function. The summation layer consists of a group of computational units with the number equal to the total number of classes.

Then finally, the output neuron(s), which is a threshold discriminator, selects the class that contains the largest response to the inputs (Ertheridge and Sviram, 1997; and Yang *et al* 1999).

The results of the logistic regression showed the important factors that can be used to determine the creditworthiness of a borrower as capital turnover ratio, total asset value and the duration of a bank-borrower relationship. It also demonstrated that credit worthiness was determined by a higher value of assets which would translate to a higher probability of a good loan. In terms of accuracy, the ANN model might not always predict default risk and the borrower's creditworthiness better than the logistic regression model. Although ANN can detect Type I error (the costs of classifying a bad loan as a good loan) it is much better than the logistic regression models

This study concluded that the use of the PNN model was better than ANN model in classifying and screening agricultural loan applications in Thailand, since it has the lowest misclassification cost.

2.8.2 The Case of a South African Credit Provider

This research aimed at determining the different factors that are used by a credit institution to evaluate loan applications in the agricultural sector. The research explored the factors used by financial institutions (Henning and Jordaan 2016) in granting credit to applicants. Some of the factors explored to see whether they had an impact on the final decision were: account standings, years as client, credit record, number of enterprise diversification, age, collateral, loan amount, education, payback period and financial performance amongst others.

Loan applications from 128 predominantly commercial farmers, were obtained from a credit institution that has several branches across South Africa. The information provided was coded by an executive representative of the financial institution so that the research team could not have any clue about the client.

The information obtained from the financial institution was guided by the research from Henning and Jordaan (2016 and 2015) and this included: credit history, purpose of the loan, collateral, amount, years as client, period of repayment, account standing, financial information, farm diversification (number of enterprises on farm), and industry risk association. Information about the applicant included: ownership of business, years of farming experience, age and education background. The final decision of the credit provider was used as a binary dependent variable, where 1 meant that the application has been approved or 0 when rejected.

Variables were categorized into three groups; financial characteristics, loan characteristics, and personal characteristics. Logistic regression was then used to investigate how the loan application variables influence the outcome of the loan application.

The logistic regression is shown in the equation below:

$$\log\left[\frac{p_i}{1-p_i}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_n X_n + \varepsilon$$

p_i is the probability of the default of an agricultural borrower i, and β_0 is the intercept term.

B_i is the respective co-efficient in the linear combination of independent variables X_i for i = 1 - n, which includes borrower characteristics, financial ratios, and farm characteristics. The dependent variable is the logarithm of the odds, $\log\left[\frac{p_i}{1-p_i}\right]$ which is the logarithm of the ratio of two probabilities of the outcome of interest (Lee *et al*, 2002). Given the set of independent variables, the probability of a value of one (1) for the dichotomous outcome is shown below.

$$\frac{p_i}{1-p_i} = \frac{1}{1+e^{-z}}$$

Where,

$$Z = \beta_0 + \beta_1 X_i + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

The objective of a logistic regression in credit scoring is to determine the conditional probability of a specific observation within a class, given the values of the independent variables of the credit applicant (Lee *et al*, 2002).

Results indicated that loan applications that were more likely to be successful were from older and more experienced farmers, with sufficient collateral, more years of business with the credit provider, request smaller loan amounts, higher production cost ratios, have lower interest expense ratio, and have diversification strategies.

The study concluded that applicants must trade-of the loan amount applied for against the size of the loan amount needed given the current debt structure and repayment ability to ensure affordability of the loan over the loan duration.

2.9 Literature review conclusion

In conclusion, the section on literature review has explored the historical development of credit scoring, its emergence, use and application. Credit scoring uses the borrower's historical data and credit characteristics to determine the borrower's creditworthiness. The literature review revealed that financial exclusion significantly impacts small holder farmers. Financial institutions require credit history to evaluate the risk of a potential borrower before making lending decisions. Most farmers face hindrances when they want to access credit facilities to fund their farming activities. In retrospect financial institutions are faced with a challenge of collecting data from farmers in far flung areas. To close this gap and reduce farmer exclusion, there was need to investigate how non-financial data which is not related to a person's financial activities can be used in credit scoring. This can be achieved by use of geospatial technology, especially remote sensing which can be applied to assess the health of crops using NDVI measurements. NDVI measurements will provide crop health assessments and even estimates of the expected yields at harvest; hence farmer credit assessment can be done objectively.

The literature review also highlighted that in Kenya, there are three main actors involved in credit scoring; insurance companies, commercial banks, and the Agricultural Finance Corporation (AFC). Credit scoring analyzes electronically the borrower's credit history and other characteristics regarding repayment ability that are, in general, provided by borrowers. It emerged also that access to credit facilities by small and poor rural farmers has a two-fold advantage; the potential of making the difference between poverty and economically secured life and can enhance agricultural productivity of farmers. Therefore, limited availability of credit can undermine income activities for farmers due to lack of capital for investment and can prevent farmers from adopting improved farming practices.

3. METHODOLOGY

3.1 The study area

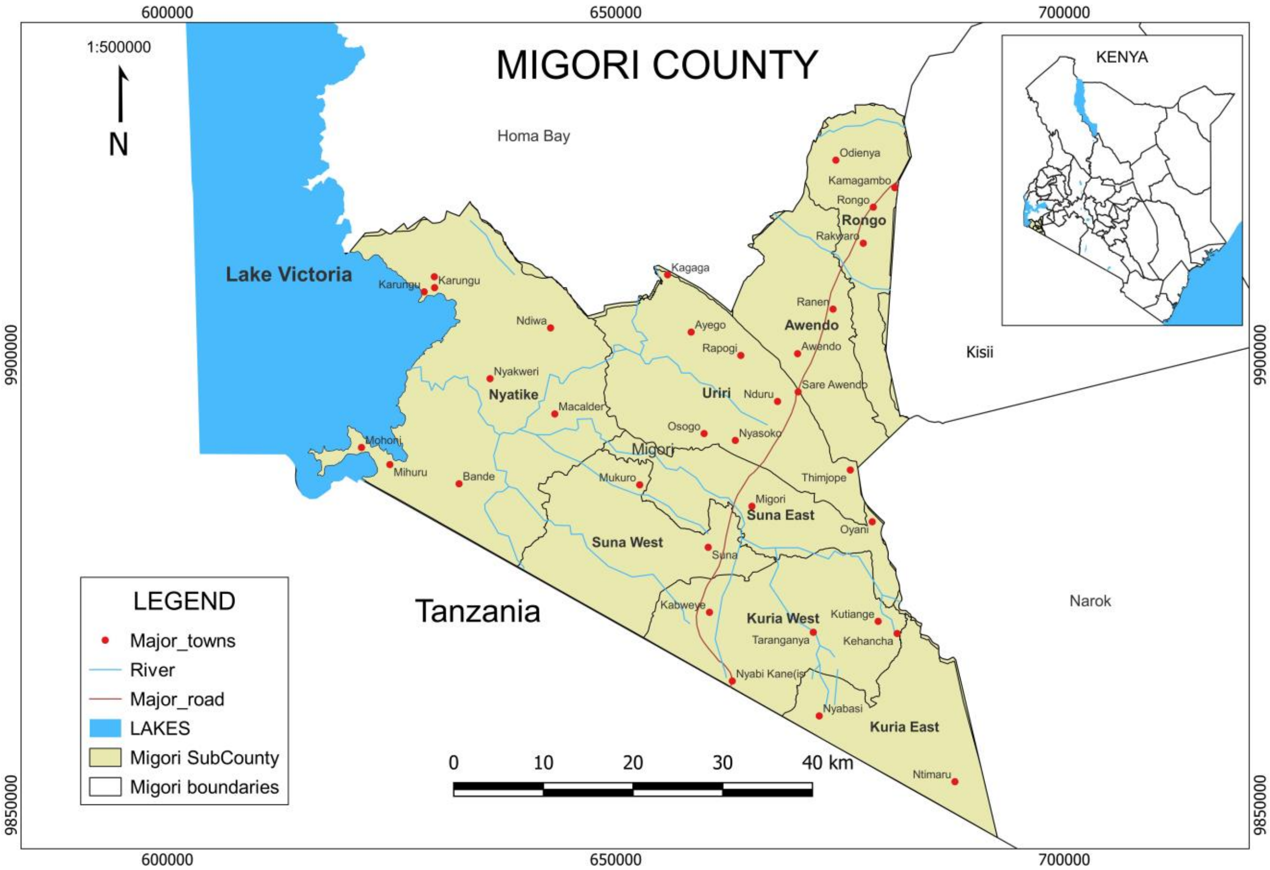

Figure 3: Map showing the study area

Migori is a county in the former Nyanza Province located in south-western Kenya and bordering the Counties of Homabay, Narok, Kisii, Tanzania and Lake Victoria. Geographically, it lies between east longitudes 33° 55' 42'' and 34° 43' 50'' and latitudes -1° 39' 06'' and -1° 23' 21'' It has eight sub-counties, namely; Awendo, Uriri, Rongo, Kuria east, Kuria west, Suna east, Suna west and Nyatike. The main livelihood activities in Migori County include agro farming and pastoral farming. Among these livelihood activities, crop farming forms the backbone of the economy in the county. At least 70% of the people residents in Migori County are dependent on crop farming, and the farming is dependent on rain fed agriculture. It is estimated that there are 960,000 small scale farmers in Migori.

Demographically, Migori County is the most diverse region of former Nyanza province after Kisumu County. The main inhabitants are Suba people, Luo, Abakuria, Abagusii, and Abaluhya. Others are Somalis, Indians, Arabs, and Nubians. The population of Migori County according to the 2019 population census was 1.2 million people (Kenya National Bureau of Statistics, 2019).

Climate wise, Migori County has two main rainy seasons. The first rainy season starts in March and ends in May, this season constitutes the long rains. The second rainy season of short rains starts in September and ends in November. The driest months are between December and February and June and September. The average daily temperature is usually a low of 24 degrees Celsius (74 F) and a high of 31 degrees Celsius (87 F). The rains often come in the afternoon and the heat is often dry and thus bearable.

Migori County was chosen for this study due its ease of access for field research, and also its prevalence of small scale maize farmers who actually do not own the land that they farm on.

3.2 Methodology: Flow chart

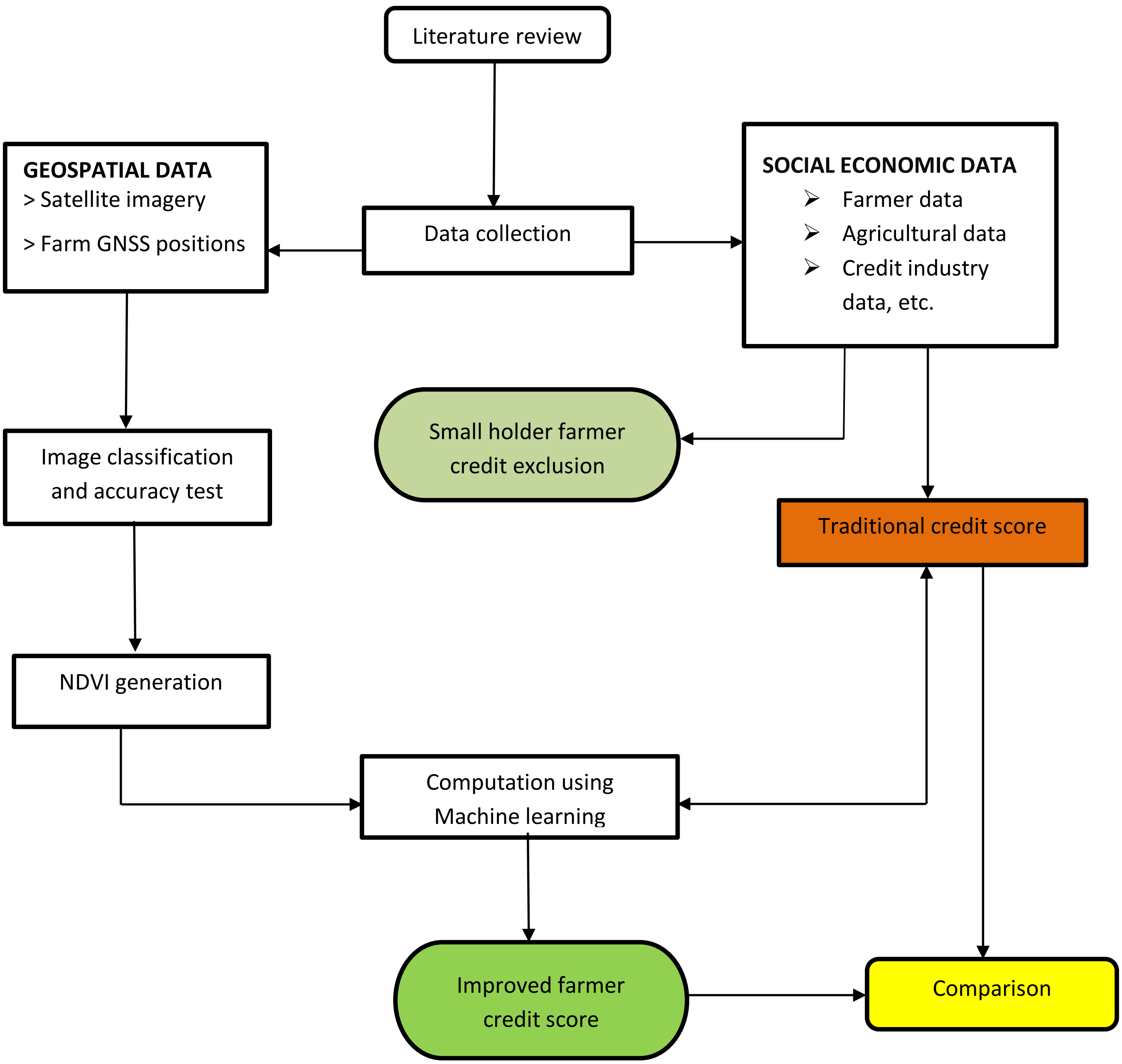

Figure 4: Methodology flow chart

3.3 Data Collection

Data collection was in two phases; phase 1 involved the administration of questionnaires to diverse respondents in the study area, while phase 2 involved collection of crop data from imagery and also from the ground by GNSS.

3.3.1 Phase 1

Small scale farmers were identified through stratified random sampling and relevant data collected from them, through a field questionnaire survey conducted in all the eight sub counties of Migori County. The data collected from each respondent included personal information (such as name, identification, gender, presence or absence of bank account, etc.), occupation, any assets, types of crops grown and access to credit. A sample of the questionnaire is presented in appendix 1. 337 questionnaires were administered with 320 being completed and returned, giving a very good return rate of 95%. Interviews with key informants in the financial institutions and Credit reference bureaus in Kenya were also conducted, in order to get a feel of their awareness of and use of credit scoring.

3.3.2 Phase 2

In this phase, Sentinel 2 imagery covering the study area was collected and LULC classified using Google Earth Engine. The date of the imagery was November 2021. In addition, reference points for later use in image analysis were selected in a well distributed pattern over the study area and positioned using hand held GNSS. Phase 2 data collection started in Kuria sub County, due to the rains and the poor road infrastructure there. 158 reference points were collected from seven sub Counties; Nyatike was excluded due to minimal farming there. To ensure even distribution of the points, it was decided to pick 5 points in each sub County, representing west, east, north, south and central areas of the sub County; for example, in Suna East sub County, points were picked from Rabuor in the east, Nyarongi in the west, Godjope in the north, Witharaga in the south and Ngege in central. Two GNSS instruments were used during the reference data collection, a Trimble TDC100 mobile mapper and a Garmin Trex 10 hand held. A research assistant was trained to operate the second instrument for the purpose of validation and back up of the data from the first instrument which was operated by the researcher herself. The collected mosaicked imagery is shown in Figure 5 and the distribution of the reference points in shown in Figure 6

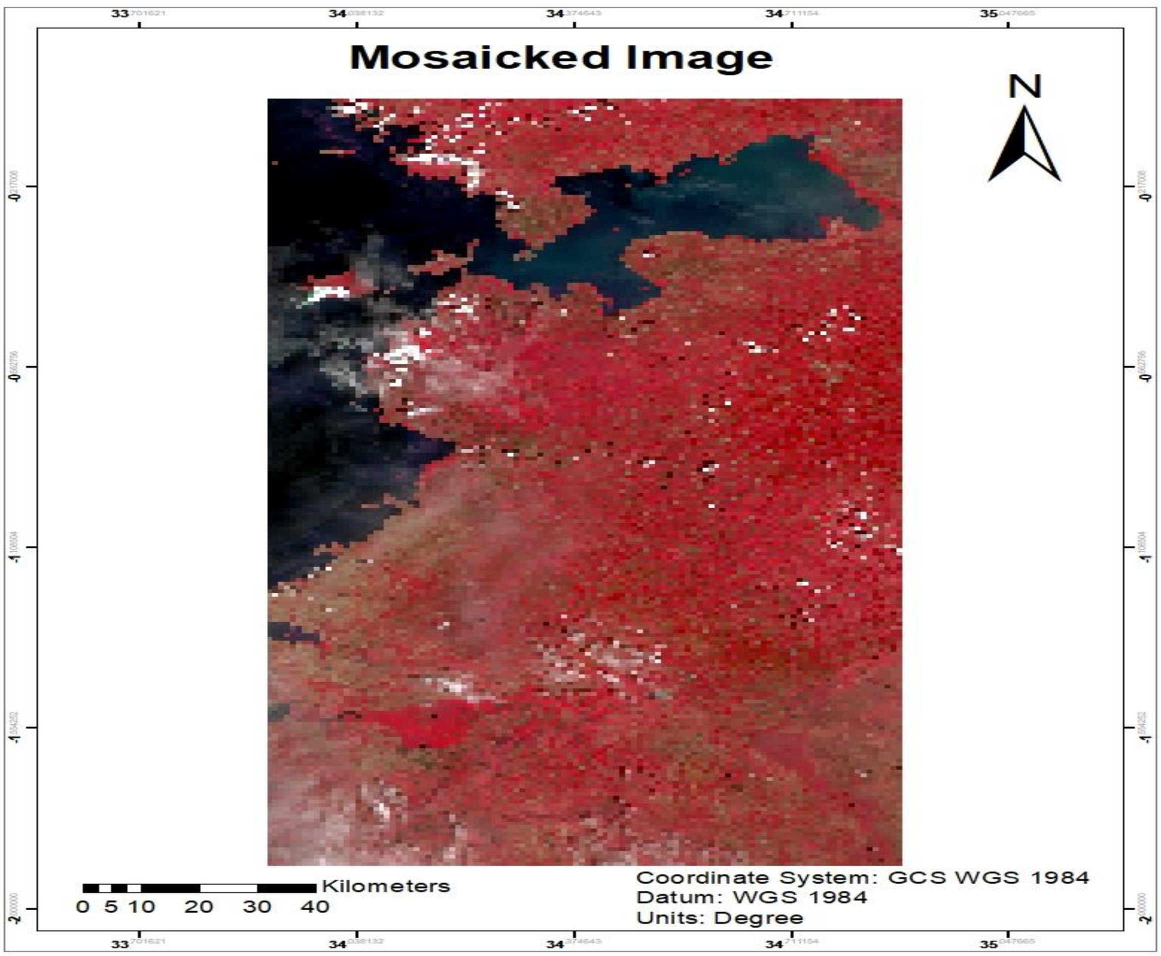

Figure 5: The mosaicked image

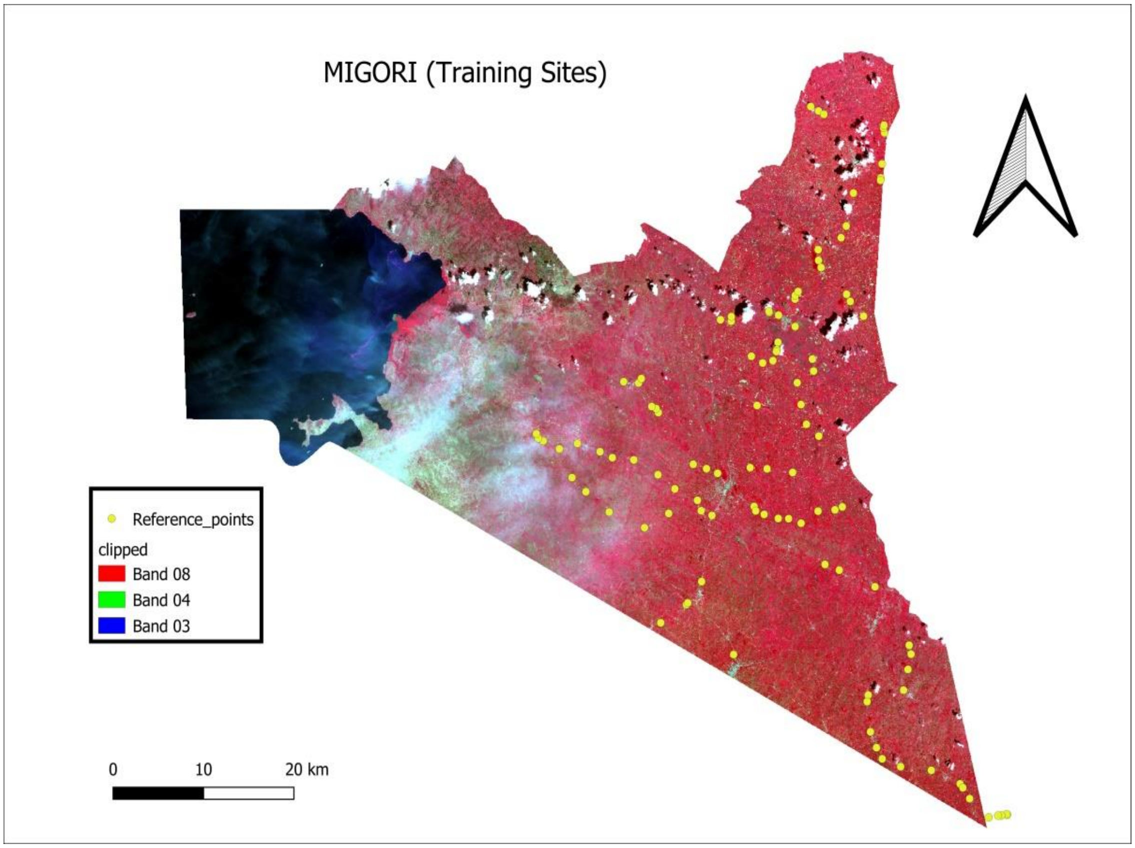

Figure 6: Distribution of the reference points

NB: (each reference point is within a candidate farm)

3.3.3: Data Preparation/ cleaning

Data collected from the questionnaires was abstracted and stored using SPSS software. Data cleaning was then done by eliminating outliers and correcting all the errors within the dataset. Other activities that were carried out during data preparation included tabulation, coding and data adjustments.

For the imagery, data preparation involved the preparation of a false color composite (8 4 3) from the downloaded data. For the ground truthing data (reference points) data preparation involved exporting the checked field data from the GNSS instruments to a laptop, and also creating a backup.

3.4: Data Analysis

The data analysis is presented as per the specific objectives.

3.4.1 Data analysis – Objective 1

This objective required the identification of the small holder farms and farmers that would be later involved in the credit scoring analysis. Candidate farms and farmers were subsequently selected using the following criteria.

- Farms size was to be less or equal to 2 hectares
- The farmers had to either have title to their farms or be farming on family land.
- For even distribution across Migori, 15 farms per sub-county were picked except for Nyatike sub-county where farming activity is minimal. In total, 101 farms were selected.
- The farms were to be in a cloud-free area in the imagery (as indicated by the corresponding reference data), to enable later generation of NDVI.

3.4.2 Data analysis – Objective 2

This objective required a determination as to whether the smallholder farmers of Migori County were financially excluded or not, and to what general extent. This was carried out through statistical analysis as detailed below.

a) The concept of Financial Exclusion

Financial exclusion describes a situation where people are unable to access financial services such as bank accounts and credit extensions because they are deemed to be too high risk (Ozili, 2021).

Financial exclusion is caused by demand side (e.g. low income, no credit history, social status etc.) and supply side (e.g. lack of insurance, lack of regular banking facilities etc.) barriers to financial inclusion.

There are two main consequences of financial exclusion:

- Financially excluded people find it harder to raise money when they need to; hence they are unlikely to improve their financial status over time and the quality of their lives remains low in the long term.
- People who are financially excluded go on to become more socially excluded as time passes by; for example they are unable to access jobs that require bank accounts to pay them, and so they become less significant in society.

The financial exclusion of small-scale farmers is of particular significance in the developing world where it impacts negatively on the availability of life's key necessity - food.

World Bank studies show that most such farmers in Sub-Saharan Africa are only able to receive payments for their agricultural produce in cash (World Bank, 2020).

b) Statistical analysis

i) A key question

A key statistical question that needed to be answered was whether the small scale farmers of Migori are indeed financially excluded or not. Questionnaire responses from the financial institutions, and from the farmers themselves, indicated that one needed to have a bank account and collateral to qualify for credit. The following farmer variables were therefore selected for the subsequent statistical analysis.

- Gender
- Age
- Bank account (presence or absence)
- Credit
- Collateral
- Credit access

ii) Data Exploratory Analysis (EDA)

An EDA was carried out in order to determine the variation that occurred within the data variables, and also the associated correlations.

a) Variations

Variation is the tendency of the values of a variable to change from measurement to measurement (Moore and McCabe, 1993). In this case a frequency cumulative plot was employed to visualize the distribution of the dataset, as shown in Figure 7; it was generated using the R code shown in Script 1.

b) Mapping Outliers

To discover outliers in the variables selected, the R code shown in script 2 was used to generate the boxplot that is presented in Figure 8.

c) Correlations

Statistical correlation measures the extent to which two variables are linearly related (meaning they change together at a constant rate) (Alreck and Settle, 1995).

The R code shown in script 3 was used to generate the correlations shown in Figure 10.

Figure 10 indicates that there are no obvious correlations between the variables.

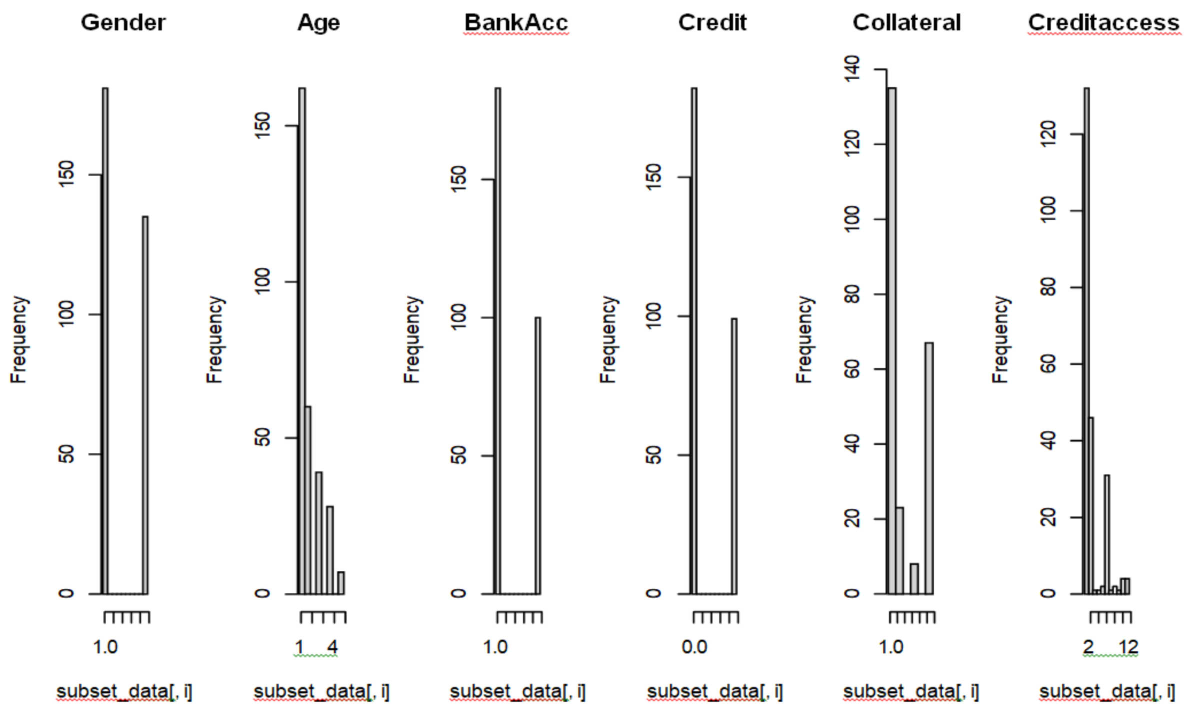

Figure 7: Frequency Cumulative plot showing how values are distributed in the given fields in the dataset

```
# The script below sets the working directory where the survey data is stored

setwd("E:/")

# Read the and store it in data variable

data = read.csv("Sue PhD.csv")

# Select the fields Gender, Age, BanlAcc, Credit, Collateral and Credit access from the dataset u
sing subset command

subset_data <- subset (data, select = c ("Gender", "Age",
          "BankAcc","Credit","Collateral","Creditaccess"))
```

To understand how data varies across different variables, a cumulative frequency was done for al
l the variables

```
# Use the par - r function to plot all the fields from 1 to 6 in the subset

par(mfrow=c(1,6))
for(I in1:6) {

hist(subset_data[,i],main=names(subset_data)[i])
}
```

Script 1: For data frequency cumulative plotting.

```
# Plot boxplot of all the variables

Par(mfrow=c(1,6))
for(Iin1:6){
boxplot(subset_data[,i],main=names(subset_data)[i])
}
```

Script 2: Data outlier mapping.

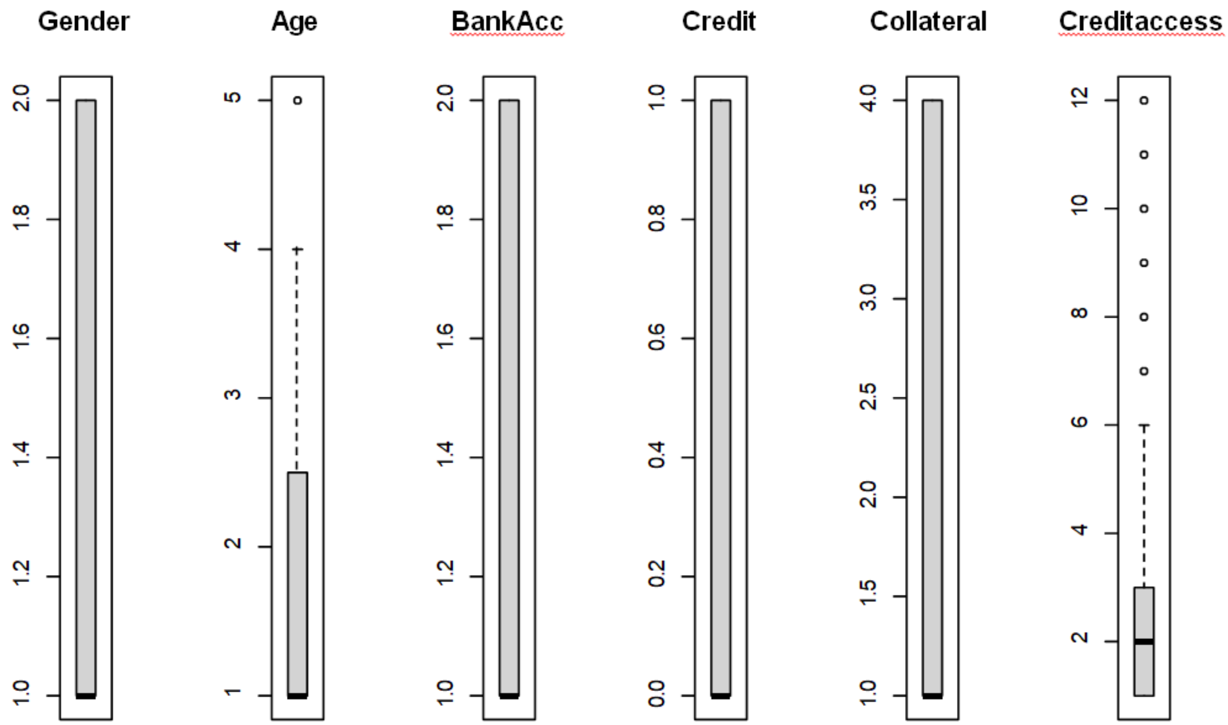

Figure 8: Box plot showing fields with outliers

```
# Load the coorplot library

library(corrplot)
correlations<-cor(subset_data[,1:6])
corrplot(correlations, method="circle")
pairs (subset_data, col=subset_data$Creditaccess)
```

Script 3: data correlation mapping

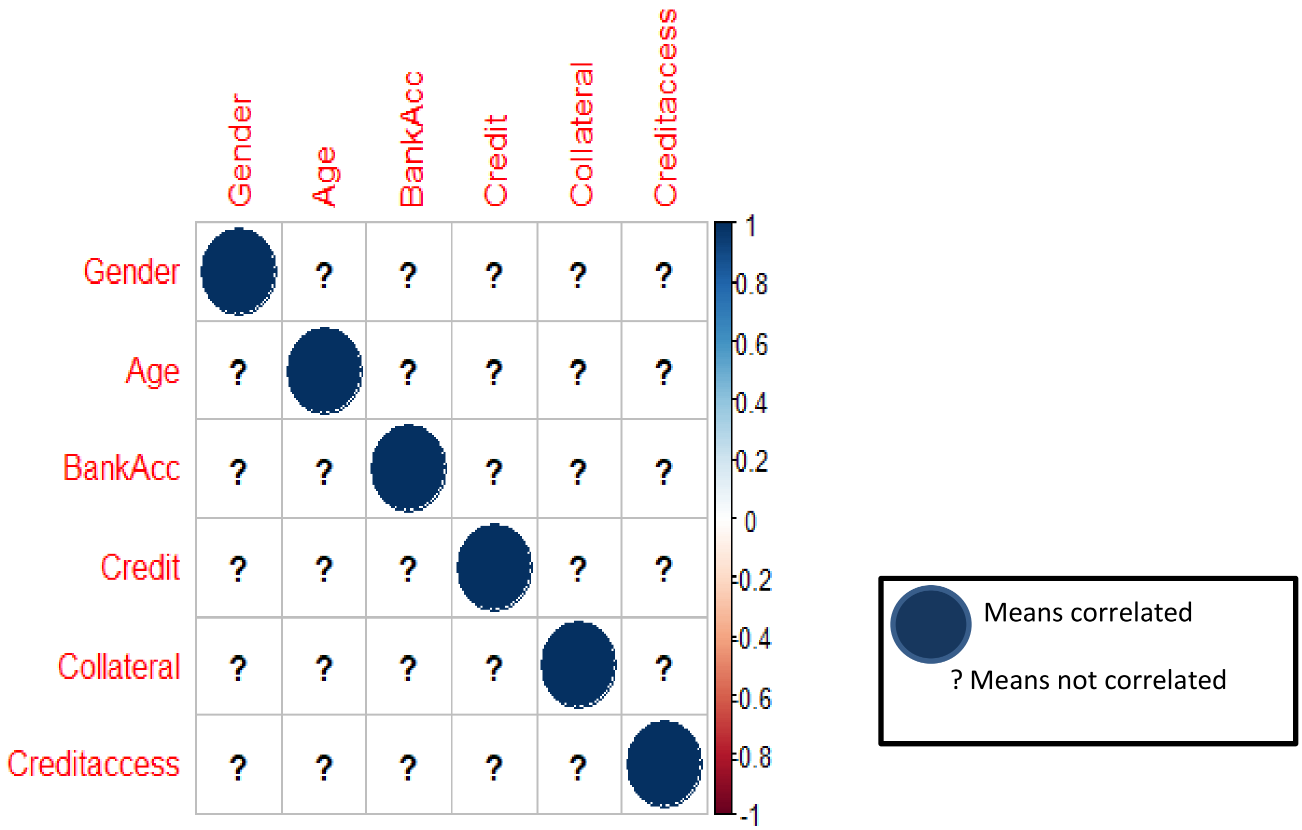

Figure 9: Data correlation map

Further examination on the density distribution of each variable broken down by credit availability value was then carried out, as per Script 4.

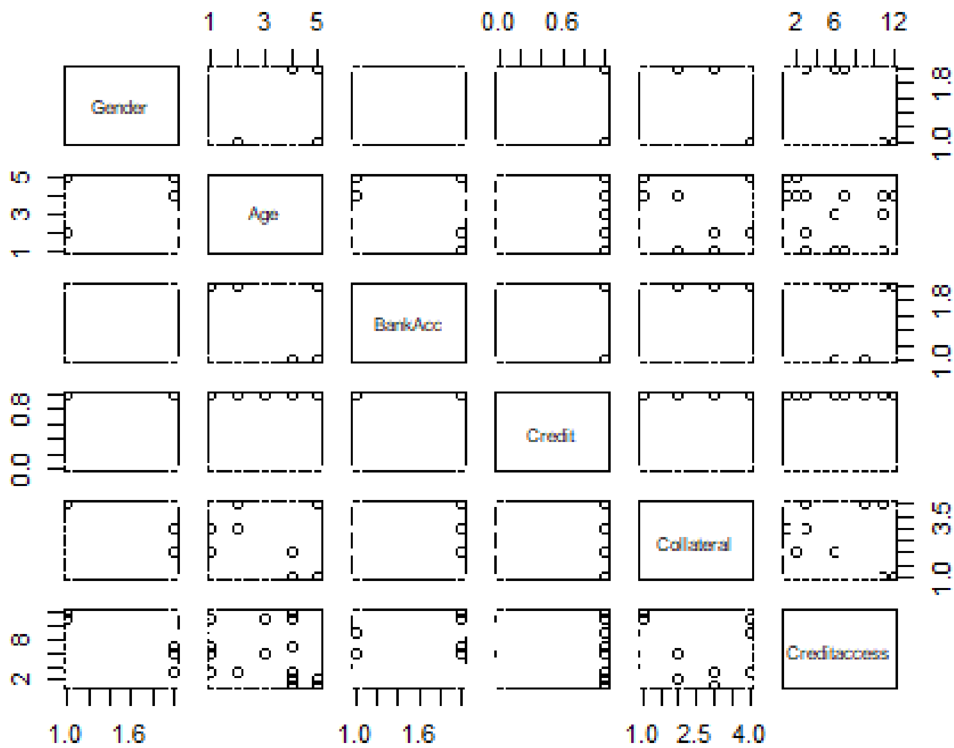

Figure 10: Variable correlations

```
# Data wrangling

library(dplyr)


subset_data$Collateral <- factor(subset_data$Collateral)

subset_data$Creditaccess <- factor(subset_data$Creditaccess)
```

Script 4: Data wrangling

iii) Modeling

A logistic regression model was used to determine whether the Migori farmers were financially excluded or not, via a hypothesis test. The logistic regression was selected due to its suitability for modeling binary variables. To model that farmers are financially included, an assumption was made that there is a relationship between being given a loan if a farmer has collateral versus being denied a loan because of lack of collateral.

Based on this Logistic regression model assumption, the null and alternative hypotheses respectively were stated as follows.

H0: $\beta0= 0$: There is no association between being granted a loan and having/lacking collateral, the odds ratio is equal to 1

H1: $\beta1= 1$ There is an association between being granted a loan and having/lacking collateral, the odds ratio is not equal to 1

The R code shown in Script 5 was used for fitting the model to the data, and the results are shown in Table 1.

From the model summary in Table 1, Collateral 4 and the intercept are statistically significant since they have a p-value of less than 0.05.

Odds Ratios were then computed at 95% Confidence Interval using the R code shown in Script 6 and the results are shown in Table 2.

Finally the model coefficients were computed using the R code shown in Script 7.

After testing the null hypothesis that there's no association between credit and collateral, the null hypothesis at the 0.05 alpha (significance) level where the z statistic is 5.327 and p-value is 9.99e−08 (refer to Table 1) is rejected.

The results show that, on average, the odds of getting credit with a group of different pieces of Collateral e.g. Title deed and Car is 7.1099742 times more than for no or one Collateral.

```
# Fit a logistic regression with credit as the response variable and
# Bank account, and Collateral as the predictors.
credit_model <- glm (Credit ~ subset_data$BankAcc + subset_data$Collateral, family = binomia
l(link = "logit"), data = subset_data)
```

Script 5: Model fitting.

Table 1: The model statistics

	Estimated	Std Error	Z value	P values
Intercept	-1.8356	0.5461	3.361	0.000776 ***
BankAcc	0.6005	0.3265	1.839	0.065933 .
Collateral2	0.4974	0.5283	0.942	0.346437
Collateral3	1.6769	0.7763	2.160	0.030756 *
Collateral4	1.9615	0.3682	5.327	9.99e-08 ***

```
# Calculate the odds ratio using the exponent function in R
exp(cbind(coef(credit_model), confint(credit_model)))
```

Script 6: Odds ratio computation.

Table 2: The model Odds ratios

		2.5%	97.5%
Intercept	0.1595139	0.05258475	0.4515562
BankAcc	1.8229853	0.96699336	3.4939391
Collateral2	1.6444898	0.55993673	4.5701397
Collateral3	5.3489107	1.20378442	28.1201725
Collateral4	7.1099742	3.52501460	15.0154637

 # Get the model coefficients

β_0=credit_model$coefficients[[1]]

β_1=credit_model$coefficients[[5]]

β_0 = -1.835624

β_1 = 1.961499

$$\ln(\text{odds of having access to credit with Collateral}) = \beta_0 + \beta_1 * (1)$$

y_0=β_0+β_1*1

= 0.1258745

$$\ln(\text{odds of having access to credit without Collateral}) = \beta_0 + \beta_1 * (0)$$

y_1=β_0+β_1*0

= 1.961499

To find the likelihood of one getting access to credit, computing of the log odds ratios as follows:

$$ln(odds(X = 1)) - ln(odds(X = 0))$$

= Y_0-Y_1

= -1.835624

$$ln(odds(X=1)) - ln(odds(X=0)) = 1.961499$$

$$e^{1.961499} = 7.1099742$$

Script 7: Model coefficient computation.

It is therefore concluded that there is sufficient evidence to reject the null hypothesis and to accept the alternative hypothesis. This means that Migori farmers cannot get credit without multiple pieces of collateral, hence they are financially excluded.

As for the extent of exclusion, the questionnaire response descriptive statistics (see Appendix 5) indicated that only 20% of the farmers have titled land under their names, which they could use as collateral. The same statistics further showed that nearly 45% of the farmers do not have any of the items that lending institutions will accept as collateral; In addition the same descriptive statistics also showed that nearly 60% of the farmers have no bank accounts. All this evidence therefore points to the fact that Migori farmers are financially excluded to a large extent. (Okeyo *et al*, 2022)

3.4.3: Data analysis – Objective 3

Data analysis in respect of this objective was in two folds; the first mainly focusing on the geospatial part which involved; image classification, accuracy testing, masking out of the maize areas and generation of NDVI. Google Earth Engine (a web portal providing global time-series satellite imagery and vector data, cloud-based computing, and access to software and algorithms for processing such data) was used to download imagery from Sentinel 2 satellite and to classify the imagery. To develop training sites, 75% of reference points earlier picked from maize farms were used as training sites for the image classification. The other classes (water, built up, bare land and other crops) were selected directly from the web.

The second part of this analysis phase focused on developing a score model that computes both the traditional score and the new scores; which include the geospatial data sets. Machine learning using logistic regression was used. This is a predominant method in credit analysis and has become the benchmark method against the other methods for such credit analysis. (listendata.com)

a) Image classification

The image was classified using the supervised classification approach and the Maximum Likelihood Classifier. The classification was based on the five classes indicated in table 3. An acceptable overall classification accuracy of 0.861 was obtained. The classified image is shown in Figure 11.

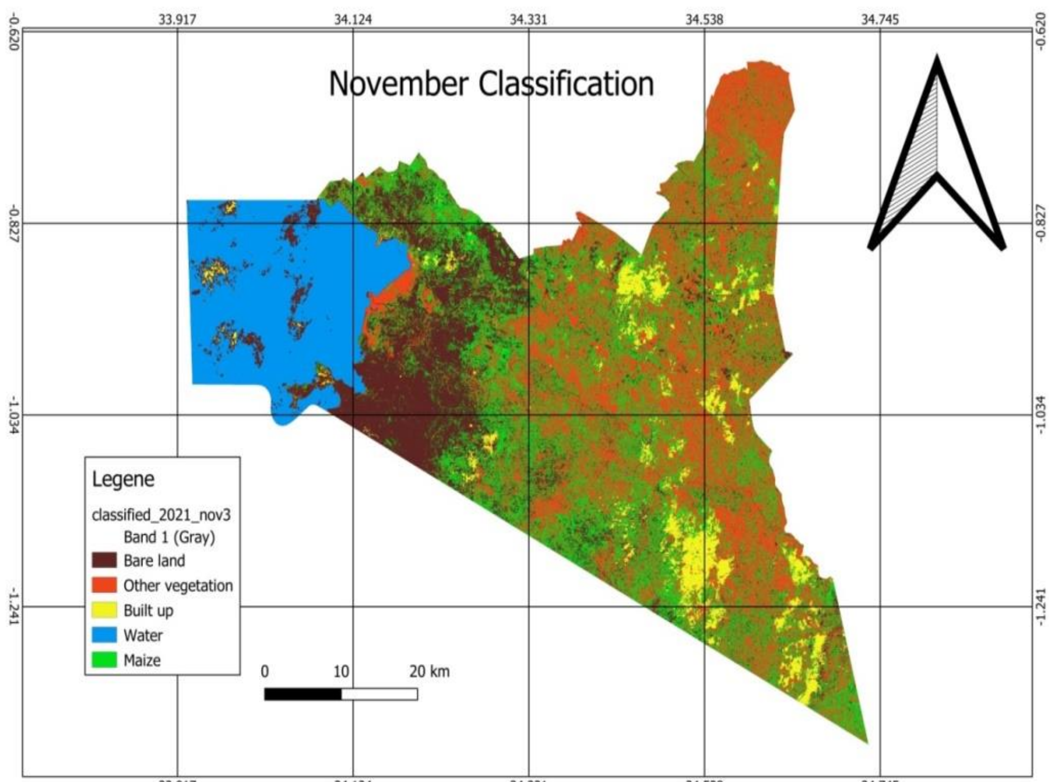

Figure 11: The classified image

Table 3: Image classification results

class	Description	Confusion matrix
2	Bare land	[0,0,30,0,0,0,1]
3	Other crops	[0,0,0,25,0,0,2]
4	Built up	[0,0,0,0,5,0,3]
5	Water	[0,0,0,0,0,3,0]
6	Maize	[0,0,2,3,0,0,5]

Classification accuracy = 0.861

b) Image masking

The classified image was now masked to only retain pixels covered by maize, as indicated in Figure 12

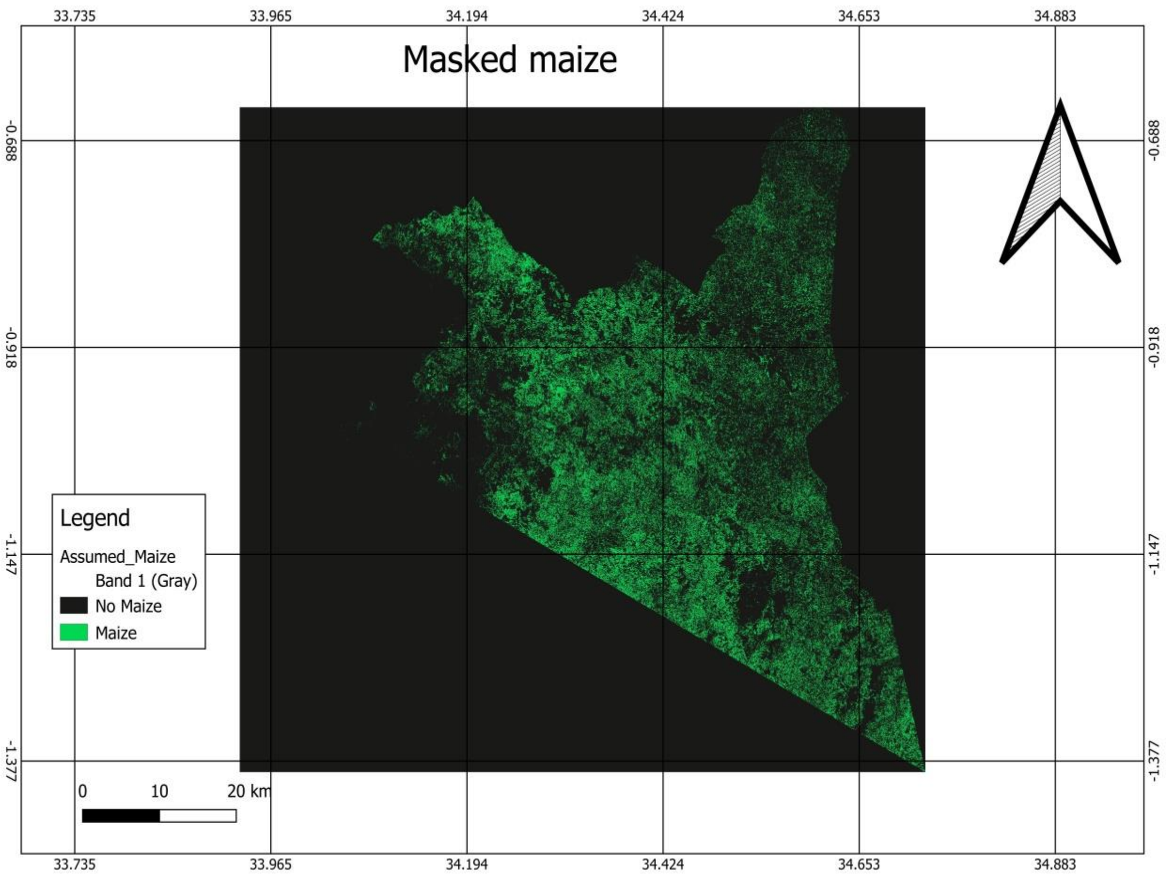

Figure 12: Masked image

c) NDVI generation and gridding

The masked maize image was now used to generate the NDVI image shown in Figure 13 using the formula: NDVI = NIR - R/NIR + R

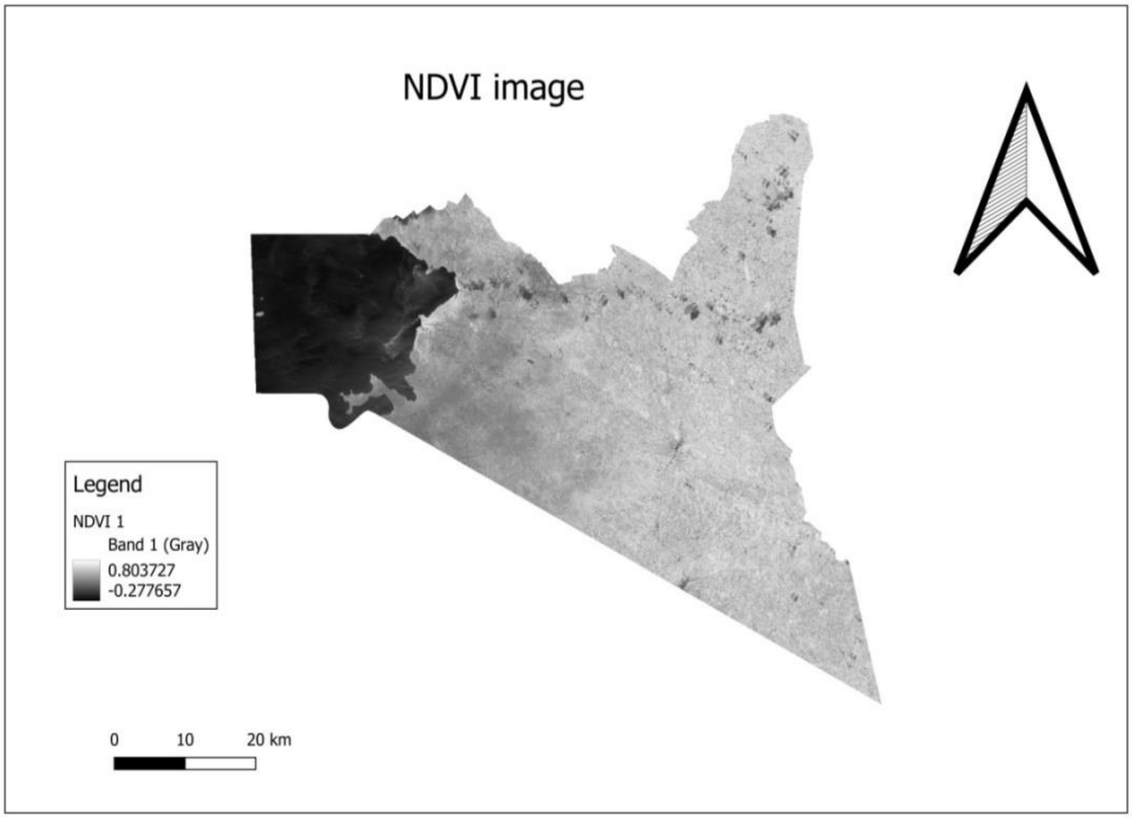

Figure 13: NDVI image

On the assumption that any small holder farm (<= 2 ha) could fit within a 1km × 1 km grid square, such a grid was overlaid onto the NDVI image and average NDVI per grid square computed. Table 4 shows an abstraction of results for NDVI averages for maize farms within the study area as identified by the reference points that had been positioned within them; the full listing is presented in Appendix 6. The whole area is about 48km × 56km.

Table 4: Abstract of NDVI averages

Farm points	Mean NDVI
Point (679440.56565792253240943 9860357.49709988757967949)	0.530378622726901
Point (679322.60980006842873991 9859664.1899594459682703)	0.509600371349184
Point (679692.43334136658813804 9856491.2468635980039835)	0.543608204119944
Point (680359.45041922363452613 9854840.06545592471957207)	0.538071751765194
Point (680988.02928306418471038 9853659.17531372234225273)	0.540530577666064
Point (683031.34388000785838813 9852836.48719156533479691)	0.51891357945354
Point (686422.71254928084090352 9852453.2730083130300045)	0.524743147588623
Point (689603.28217022749595344 9851039.77896921150386333)	0.553811495184186
Point (689916.28274229553062469 9850623.3033433835953474)	0.543480003895261
Point (690661.03552251192741096 9849483.75616908445954323)	0.546819193696806
Point (692769.64183439291082323 9847513.47706030867993832)	0.53040276924139
Point (683355.97616871655918658 9860876.84095096960663795)	0.502420804671969
Point (683837.02257441938854754 9863050.41132395341992378)	0.53192711945134
Point (684185.59619820269290358 9864655.58410394564270973)	0.514158010973312
Point (683971.26121742837131023 9865596.99114113114774227)	0.531419316457991

d) Development of the score model

This process involved modeling farmers' data to come up with a credit scoring tool to assess the eligibility of a farmer accessing a loan, with additional emphasis on crop geospatial data/information. This was based on sample farmers from Migori County who were interviewed with regards to access of credit. Weight of Evidence (WOE), a statistical technique commonly used in credit scoring to evaluate the predictive power of various features or variables was employed. This method enabled the transformation of raw data into meaningful and informative predictors, providing a solid foundation for accurate credit risk assessment. Using the R programming language, the step-by-step processes which were undertaken in implementing credit scoring using WOE and Generalized linear model are now outlined.

i) Start variables

The start variables which were obtained from the questionnaire are as shown in Table 5.

Table 5: Model start variables

Variable	**Meaning**
Gender	Gender of respondent
Age	Age of respondent
Marital	Marital status
Occupation	What farming activities does the respondent conduct?
Year of Occupation	How long has the respondent practiced farming?
How much land	How much land holding does the respondent have?
Land ownership	What ownership do you have on the land you are farming on?
Location	Where is the land located?
Land parcel	Do you have more than one parcel of land?
How many	If yes, how many parcels do you have?
Landhold	Is the land freehold or leasehold?
Market	How and where do you sell your maize product?
Collateral	What do financial institutions require as collateral before granting loan?
Creditaccess	What are challenges faced when applying for loan?
Creditrepayment	If you have previously applied for loan, did you pay back on time?
Reasons	If no specify the reasons
Acquire	Where do you acquire your loan from?
Capacity	What is your credit capacity?
How long	For how long have you been taking loans over the years?
Outstanding	Do you have outstanding loans?
Loanduration	If yes when is your loan due?
Set	How much have you offset so far?
Employment	Are you currently employed?
Employer	If yes, indicate your employer
Income	If no, indicate your source of income
Cycle	Indicate the number of crop cycle per year in your farm
Pest	What type of pest attacks do you encounter in your farming activities?
Organic	Do you use any organic pest control technique?
Specify	If yes, specify
Irrigation	Do you use any type of irrigation techniques ?
Specify irrigation	If yes, specify
Ploughing	Which ploughing methods do you use in farming?

ii) Correlation testing

All the start variables were tested for correlation. This is because the predictor variables used in logistic regression should, ideally not be correlated. Correlation among such variables can cause model problems such as multi collinearity, leading to unstable and unreliable estimate of the regression coefficients. In such cases, the regression coefficient may change dramatically with minor changes to the data. The results of the correlation testing are presented as Appendix 2, but are summarized in Figure 14, in which red color denotes positive correlation while blue color denotes negative correlation

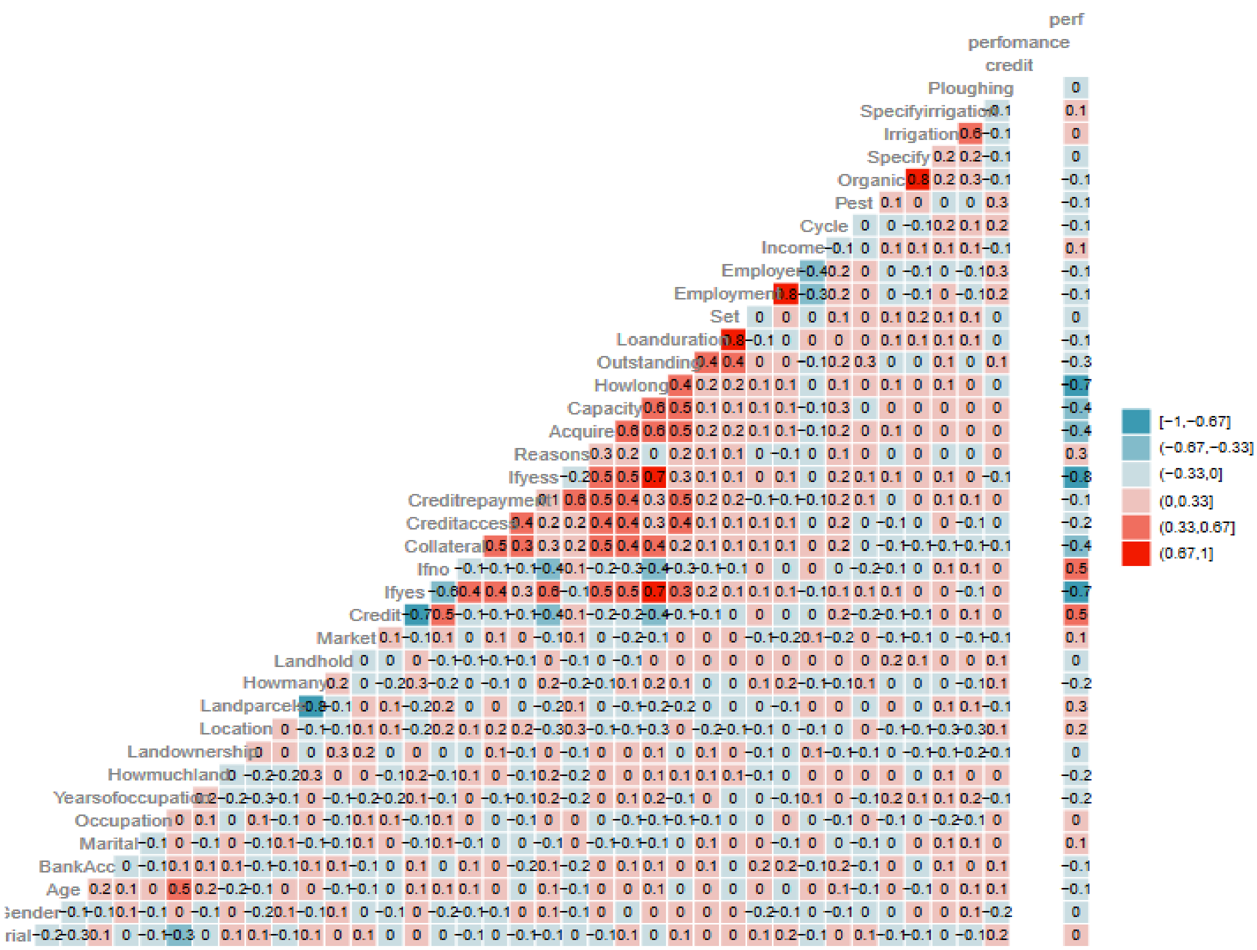

Figure 14: Correlation matrix/chart

iii) Final variable selection

Following the dropping of correlated variables, four final predictor variables were selected based on their information value (IV) and domain knowledge (i.e. knowledge of the credit scoring industry). IV is a tool in machine learning used to assess the predictive power of any variable for a given feature in a dataset, and is often used in credit scoring. It quantifies the extent to which a variable can differentiate between different outcomes, such as default and non-default, and provides insights into the variable's contribution to the predictive model's performance. IV depends on a variable's weight of evidence (WOE).

Statistically,

$$\text{WOE} = \ln\left(\frac{\%\, of\ non\text{-}events}{\%\, of\ events}\right)$$

and IV = (% of non-events - % of events)

In the case of separating loan defaulters from non-defaulters, default represents an event while non-default represents a non-event.

Generally the higher the IV, the better the variable for the intended prediction; however the best IV values range between 0.3 and 0.5; although values between 0.1 and 0.3 may also be accepted for model development as medium predictors. (listendata.com)

Figure 15 shows examples of WOE and IV results for the variables Howlong and Collateral.

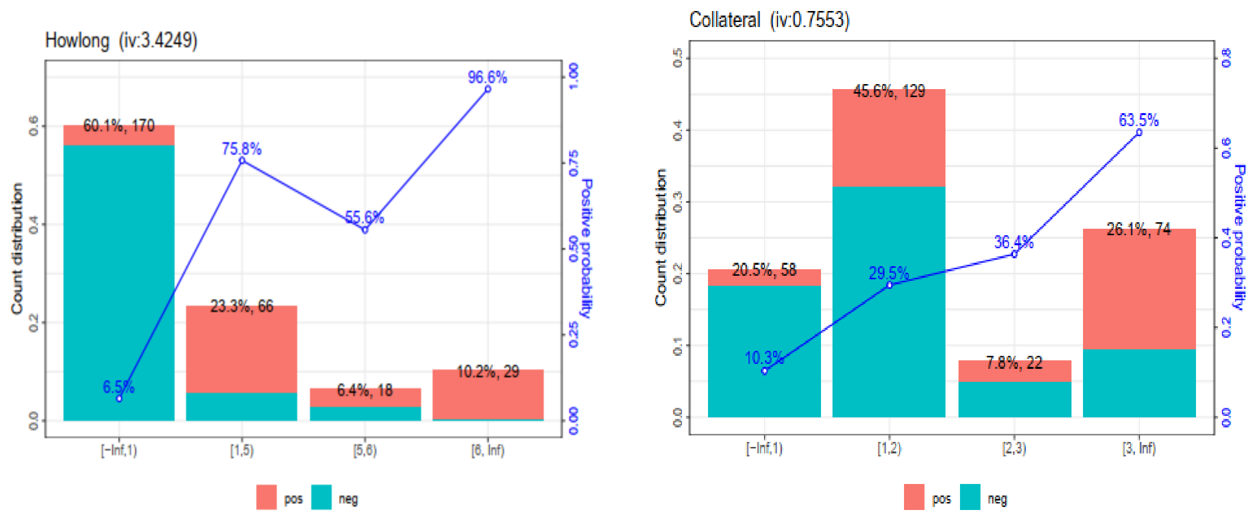

Figure 15: Plots of Howlong WOE and Collateral WOE

Following the WOE/IV analysis, the variables shown in Table 6 were selected for model development. NDVI (Normalized Difference Vegetation Index) is an important index used in remote sensing to assess and monitor vegetation health and vigour, and it is the geospatial variable that was introduced into the model to indicate whether the farmer had healthy crops or not.

Table 6: Information value of the selected variables

Variable	IV score	Remarks
Landparcels	0.273	Though low IV score the variable was selected based on domain knowledge
Collateral	0.766	
Reasons	0.137	Though low IV score the variable was selected based on domain knowledge
Howlong	2.812	
NDVI	Was not scored	Deliberately introduced

iv) Data Modelling

After converting the original data points to WOE, a Generalized Linear Model was fitted based on the following logistic regression formula.

$$\ln(\frac{P(X)}{1 - P(X)}) = intercept + B_1x + B_2x + \cdots B_nx$$

The R link function used to generate results is the logit value

$$model\_fit = glm(formula$$
$$= Credit \sim Landparcel_{woe} + Collateral_{woe} + Reasons_{woe} + Howlong_{woe}$$
$$+ NDVI_{woe}, family = binomial(link = \text{logit , data=train\_woe)}$$

Script 8: Data modelling

The output of this model, which was in logit was converted to odds and then probability as follows:

Odds = e^{logit}

Probability $P(X) = \frac{odds}{odds+1}$

After running the model, the coefficients achieved were in terms of log odds; which were converted to probabilities and then final scores (scaling). These scores were originally in the scale of 0 to 1000. The probabilities were converted to scores using the Scorecard package in R using the following approach (Script 9).

$$\text{score} = \text{offset} - \text{factor}\ \ln(\text{odds})$$

$$factor = \frac{pdo}{\ln(2)}$$

$$\text{offset} = \text{Ts} - \text{factor}\ln(\text{To})$$

Replace odds with logit

$$odds = e^{logit}$$

$$\text{score} = \text{offset} - \text{factor logit}$$

Ts – target score

To – target odds

Pdo - slopes

Script 9: Score computation

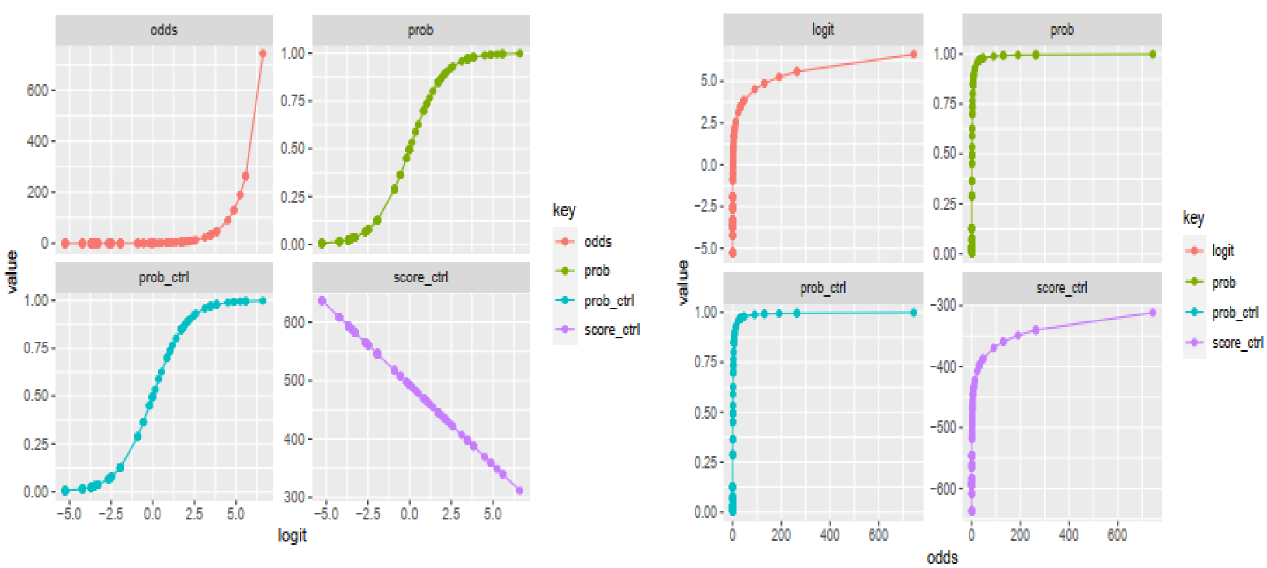

Figure 16: Logit vs. Odds, Probabilities and Score plots

The original scores were finally scaled to FICO score, which is more universal and frequently used. (myfico.com). This was done using the following formula (Shukla and Jha, 2018):

OldRange = (OldMax - OldMin)

NewRange = (NewMax - NewMin)

NewValue = (((OldValue - OldMin) * NewRange) / OldRange) + NewMin)

A FICO score is a three-digit number, typically in a 300-850 range, that tells lenders how likely a consumer is to repay borrowed money based on their credit history. Only FICO Scores are created by the Fair Isaac Corporation and are used by over 90% of top lenders when making lending decisions.

The FICO score range is explained in Table 7.

Table 7: FICO score range

FICO Score Ranges	**Credit Rating**	**Description**
<580	Poor	Your score is well below the average score of the consumers and demonstrates to lenders that you are a risky borrower.
580-669	Fair	Your score is below the average score of the consumers, though many lenders will approve loans with this score.
670-739	Good	Your score is near or slightly above the average of the consumers and most lenders consider this a good score.
740-799	Very Good	Your score is above the average of the consumers and demonstrates to lenders that you are a very dependable borrower.
800+	Exceptional	Your score is well above the average score of the consumers and clearly demonstrates to lenders that you are an exceptional borrower.

Out of the 101 farmers and farms identified from section 3.4.1, 67 (about 2/3) were chosen for training the model, with the remaining 34 reserved for later testing and validation of the model.

v) Model Evaluation

The performance of a binary classification model, such as the logistic regression model used in this study, is often evaluated by plotting the Receiver Operating Characteristic (ROC) curve and determining the Area under the Curve (AUC). The curve checks how well the model is able to distinguish and separate events from non-events; it is a plot of the rate of true positives (events correctly predicted to be events) on the y-axis against the rate of false positives (non-events wrongly predicted to the events) on the x-axis. Generally the higher the AUC curve i.e the bigger the area under the cover, the better the model, and 75% is the recommended acceptable minimum. Script 10 was used to determine this AUC. A related evaluation measure is the Gini coefficient defined as: Gini = (2 * AUC – 1).

Again, a higher Gini represents a better predictive model.

```
## A performance instance
##   'Area under the ROC curve'
```

####Make prediction

Testing of the predictive power of the model on test data that had been earlier isolated is done.

```
## [1] "K-S Statistic = 0.846718005133847"

## [1] "Area under the curve = 0.961037770443711"

## [1] "Gini Coefficient = 0.922075540887423"

##     (Intercept)     Landparcels       Landhold      Collateral Creditrepayment
##   -5.2637155148   -1.5550051782   -0.5318544217    0.4488074659   -0.6338934531
##        Reasons         Acquire         Howlong     Outstanding             Set
```

```
##   -2.6601538908    0.6818583697    0.3192265714    3.1808766951   -0.0006021895
##           Pest         Organic
##    0.0635936312    1.5129961303
```

Script 10: Evaluating the model by AUC

4.0: RESULTS and ANALYSIS OF RESULTS

4.1: Results

a) Objective 1

This objective required the identification of the small holder farms and farmers that would be later involved in the credit scoring analysis. From the methodology in section 3.4.1, 101 farms and farmers were identified, and Table 8 shows an abstract. The full list is in Appendix 3.

Table 8: Abstract of small holder farms and farmers

SERIAL NO	FARM ID	FARM LAT	FARM LONG	FARMER ID	FARM NDVI AVERAGE	Sub-county
1	17	-0.772	34.625	112548	0.58	Suna East
2	4	-0.88	34.543	221586	0.55	Rongo
3	102	-0.724	34.628	245812	0.60	Suna East
4	10	-0.937	34.521	362547	0.33	Suna West
5	42	-1.084	34.499	1013025	0.56	Rongo
6	60	-0.988	34.502	1076489	0.54	Rongo
7	68	-0.945	34.518	1121536	0.56	Rongo
8	216	-1.258	34.648	1847119	0.50	Kuria West
9	43	-1.088	34.501	2325116	0.58	Rongo
10	50	-1.084	34.587	2403160	0.53	Rongo

b) Objective 2

This objective required a determination as to whether the small holder farmers of Migori County were financially excluded or not, and to what general extent. From the statistical analysis set out in section 3.4.2, it was found that indeed these small holder farmers are financially excluded to a large extent.

c) Objective 3

This objective required the generation of average NDVI values per participating farm, then development of a credit scoring model. From section 3.4.3, it can be seen that the study area color composite was LULC classified with a good classification accuracy of 86% (Figure 11, Table 3), area covered with maize isolated (Figure 12) and average NDVI per farm computed (Table 4).

On model development, the final model developed is shown in Table 9, with the indicated predictor variables; the more the stars, the more statistically significant the variable. However, statistical significance may be different from practical significance.

Table 9: Credit scoring model statistics

	Estimates	Std. Error	z value	Pr(>|z|)	Significance
Intercept	-1.2873	0.3705	-3.474	0.000512	***
Landparcels_woe	1.0331	0.5107	2.023	0.043081	*
Collateral_woe	0.8495	0.3020	2.813	0.004903	**
Reasons_woe	1.4431	0.4171	3.460	0.000540	***
Howlong_woe	1.0761	0.1874	5.742	9.38e-09	***
NDVI_woe	2.0734	0.8762	2.366	0.017967	*

The model was trained on 67 farms and farmers, and the results are abstracted in Table 10 and fully presented in Appendix 4.

Table 10: Abstract of model training results

Farm serial	Farm id	Farmer ID	Land Parcels	Collateral	Reasons	Howlong	NDVI	Previous Credit	Traditional score	With 1.0 NDVI	With 1.2 NDVI	With 1.3 NDVI	With 1.4 NDVI	With 1.5 NDVI
1	1	20184113	2	1	Null	Null	0.64	0	685	722	806	848	850	850
2	2	27186633	2	5	Null	Null	0.62	0	599	641	709	743	777	811
3	3	8144389	1	1	Null	Null	0.63	0	639	663	735	771	807	844
4	4	221586	2	1	Null	Null	0.62	0	685	722	806	848	850	850
5	5	32261560	2	Null	Null	Null	0.56	0	726	661	733	769	805	841
6	7	2803912	2	4	Null	20	0.52	1	367	408	430	441	452	463
7	8	37703733	2	2	Null	1	0.30	0	510	499	538	558	578	598
8	10	362547	2	Null	Null	Null	0.60	0	726	661	733	769	805	841
9	11	30142286	2	Null	Null	5	0.33	0	565	549	598	623	648	673
10	13	22211547	2	Null	Null	Null	0.58	0	726	661	733	769	805	841

In respect of Table 10, it should be noted that the model was first run **without** NDVI, to generate the 'traditional' score column. NDVI was then introduced, to generate the 'with 1.0 NDVI' column.

For experimental purposes and in view of the central role that NDVI was to play in this whole arrangement, the weight of NDVI was deliberately biased to 1.2, 1.3, 1.4 and 1.5 the original weight and this generated the correspondingly labeled columns.

The performance of the model in this training was evaluated by plotting the ROC curve and determining the AUC; the resultant curve is shown in Figure 17.

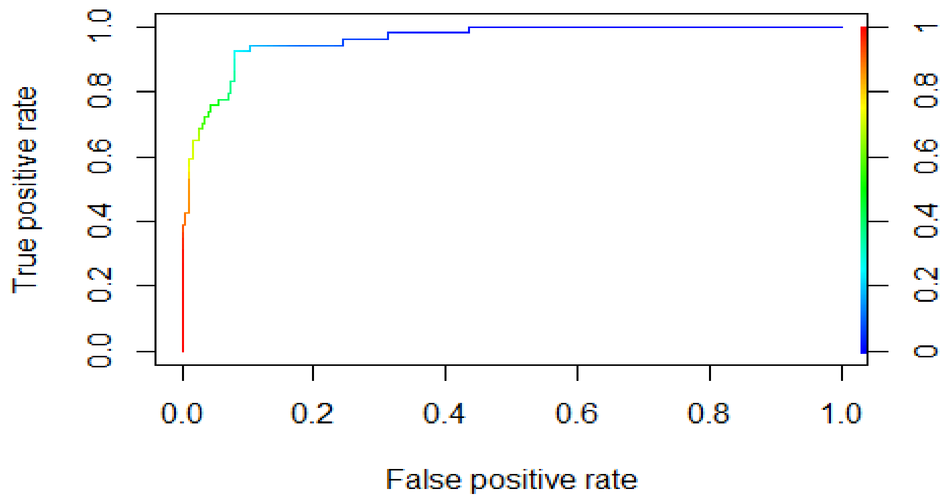

Figure 17: AUC curve

The color in the AUC curve corresponds to the area under the curve, with red representing the least and blue the highest.

From the curve, the AUC was found to be 95%, with a Gini Coefficient of 0.9.

Following these encouraging results from model training, the model was now tested on the remaining 34 farms and farmers, and the full results are shown in Table 11.

Table 11: Model testing results

Farm Serial	Farm Id	Farmer ID	Land Parcels >1?	Collateral	Reasons	Howlong	Previous Credit	NDVI	Traditional Score	With 1.0 NDVI	With 1.2 NDVI	With 1.3 NDVI	With 1.4 NDVI	With 1.5 NDVI
1	6	13388005	2	Null	Null	Null	0	0.74	726	760	850	850	850	850
2	9	6912420	2	1	2	Null	0	0.44	825	823	850	850	850	850
3	12	13112933	2	Null	Null	2	0	0.67	562	585	642	670	699	727
4	14	21186481	2	Null	Null	Null	0	0.56	726	661	733	850	805	841
5	16	21086953	2	Null	Null	3	0	0.57	515	434	461	580	488	501
6	18	24244325	2	2	2	5	0	0.58	635	564	616	735	669	695
7	19	3369874	2	Null	2	Null	0	0.61	850	850	850	850	850	850
8	39	3654821	2	1	Null	4	1	0.46	474	511	553	526	596	617
9	40	8159461	1	Null	Null	Null	0	0.51	681	717	800	795	850	850
10	42	10130257	2	Null	Null	Null	0	0.69	726	760	850	850	850	850
11	50	24031606	1	5	Null	4	1	0.60	342	354	365	354	376	382
12	51	9771763	2	Null	Null	Null	0	0.61	726	760	850	850	850	850
13	52	33724440	2	Null	2	2	0	0.62	703	726	811	823	850	850
14	56	5643871	2	Null	Null	Null	0	0.45	726	721	805	850	850	850
15	57	27993343	2	Null	Null	Null	0	0.38	726	721	805	850	850	850
16	60	24922360	2	1	Null	Null	0	0.49	685	738	825	801	850	850
17	63	23890276	1	8	Null	10	1	0.64	360	369	383	378	396	403
18	70	22407658	2	4	2	10	1	0.40	507	494	532	570	571	590
19	71	22304385	1	1	Null	16	1	0.63	408	414	437	440	460	472
20	72	24503941	2	1	Null	Null	0	0.59	685	622	687	801	751	784
21	75	37752062	1	1	Null	3	1	0.45	428	396	416	467	435	444
22	125	11215846	2	6	Null	4	1	0.57	388	314	317	414	320	322
23	126	30002630	2	1	Null	5	0	0.65	524	550	600	591	650	675
24	127	6912045	2	1	Null	7	0	0.48	524	566	619	591	672	699
25	129	3972355	1	10	Null	2	1	0.54	428	342	350	466	359	363
26	132	24222361	1	1	Null	5	0	0.26	478	451	481	532	511	526

27	134	26811913	2	1	Null	3	1	0.56	474	396	415	526	434	444
28	136	32241069	2	1	Null	Null	0	0.58	685	622	687	801	751	784
29	145	29547139	2	1	Null	Null	0	0.52	685	738	825	801	850	850
30	204	11224046	2	2	Null	Null	0	0.62	656	694	773	763	850	850
31	210	7.21E+08	1	1	Null	Null	0	0.54	639	563	616	741	668	694
32	217	25696211	2	7	2	Null	0	0.50	777	833	850	850	850	850
33	221	11412683	2	1	2	Null	0	0.62	825	850	850	850	850	850
34	222	13244463	2	Null	8	Null	0	0.31	850	850	850	850	850	850

4.2: Analysis of Results

a) Objective 1

The results for this objective showed that 101 farmers and farms were selected, distributed as follows; 39 from Rongo, 14 from Suna East, 21 from Suna West, 6 from Awendo, 6 from Kuria East, 11 from Kuria West and 4 from Uriri. Most farms from Uriri and Awendo were not selected because most farmers from these areas were farming on rented lands and therefore did not qualify to be selected for sampling. On the other hand, most farms in Rongo qualified to be sampled as the farmers had titles to their land or were farming on family land. Notably, most farmers around Kuria East and Kuria West had farms greater than 2 ha, hence this did not also qualify them for sampling. The splitting of the 101 into 67 for training and 34 for testing the model (i.e. a 2:1 ratio) followed a pattern observed in some previous studies, such as Simumba (2018).

b) Objective 2

The results for this objective showed that indeed, these small holder farmers are financially excluded to a large extent. This is illustrated in the statistical analysis set out in Section 3.4.2; also from the data descriptive analysis in Appendix 5, it was revealed that only 32% of the respondent had bank accounts. Clearly therefore, any intervention that could improve the creditworthiness of small scale farmers would be beneficial to the farmers of Migori and similar counties.

c) Objective 3

The credibility of the scoring model developed is evidenced by the ROC curve in Figure 17, whose high AUC indicates a good choice of predictors. The results in the Table 11 show that out of 34 farmers scored, 14 were poorly scored (below 580) and hence would not be recommended for credit. On introduction of the NDVI variable, nearly half of all farmers had their scores improved, and one farmer was able to move to a score of over 580 and hence become eligible for credit. On enhancing the weight of NDVI up to 1.5 NDVI, 5 more farmers were able to transition to eligible status. Therefore by the end of the NDVI experiment, 6 farmers out of the originally ineligible 14 had changed status and could now get credit; the level of non-eligibility had therefore changed from 41% to 24%. This is a promising result that can be built upon if the lending industry were to warm up to it.

5.0: IMPLEMENTATION FRAMEWORK

" It is always easier to talk about change than to make it…Even the best strategies fail to take into account more than a few of the consequences that flow from them…" (Toffler, 1985).

It has been demonstrated in Chapter 4 that the introduction of NDVI as a factor in the evaluation of Migori small holder farmers for credit holds the promise of improving their credit scores, hence their potential to access credit. Considering that most lending institutions now provide credit based almost entirely on the financial data of the applicants, this new idea (of considering NDVI) could excite both the small scale farmers and the lending institutions; for the former it holds the promise of redeeming them from long standing financial exclusion, and for the latter, it holds the potential to increase their customer base by attaching a reliable attribute (i.e. a measure of crop health) to a farmer that may otherwise have very little going for them in terms of financial assets, collateral or credit history.

This chapter looks at some of the many issues that would need to be addressed in developing a program to implement this idea in Kenya.

5.1: Fine tuning the model

The model as it stands now requires further research, and the following are some suggestions towards fine tuning it through such research.

- The questionnaire survey done involved a small sample of farmers; better results could be obtained from a larger sample. This would improve on the analysis of the data process by eliminating and reducing extreme values or outliers on scores generated that usually result from small data samples. Generally, the more data points there are, the better one can train a machine learning model.

- The questionnaire administered should also focus on personal detail such as collateral that each individual farmer has instead of generalizing on the collateral required by the financial institutions.
- The geospatial component of the model could be improved by adding to NDVI data such as rainfall, soil properties, evapotranspiration etc.
- To be a viable idea for the whole country, the fine-tuned model would need to be tested on small scale farmers from more counties across Kenyan agro-climatic landscape.

5.2: Implementation issues

5.2.1: Policy issues

Lending is a program that has been run over the years by financial institutions, and the relevant policies vary from country to country. In Kenya, the government implements such programs through the Central Bank of Kenya (CBK), which is responsible for lending and regulating on how financial institutions lend. This is usually implemented through CRBs where all banks in Kenya must report credit information on consumers (Credit Reference Bureau Regulations, 2013, sect. 18).

A key policy currently adopted is that most financial institutions rely on previous historical records to assess their borrowers' creditworthiness. A better approach would be the inclusion of predictive data, such as NDVI and others, in the system.

5.2.2: Institutional issues

Most financial institutions in Kenya use the traditional credit score model to assess farmers. This study has shown that most of these farmers are unable to access credit because of the fact that they are financially excluded. Inclusion of geospatial data to resolve this issue will require expertise to handle such a system, and also awareness of this new development, by:

- Creating awareness amongst the stakeholders such as lending institutions, Credit Reference Bureaus and the farmers themselves.
- Outsourcing or hiring of geospatial experts into the lending organization who will be in charge handling the geospatial data
- Training of staff within the organization on how geospatial data will work in their lending systems.
- Setting up a new service provider to service lending institutions and CRBs with the geospatial data that they need.

Looking at the implementation nationally, new service providers would need to liaise with related services such as the meteorological service, satellite data service, geospatial positioning service, etc. in order to provide geospatial data to the financial institutions or CRBs in order to enable the new scoring system. This is illustrated in the institutional architecture in Figure 18.

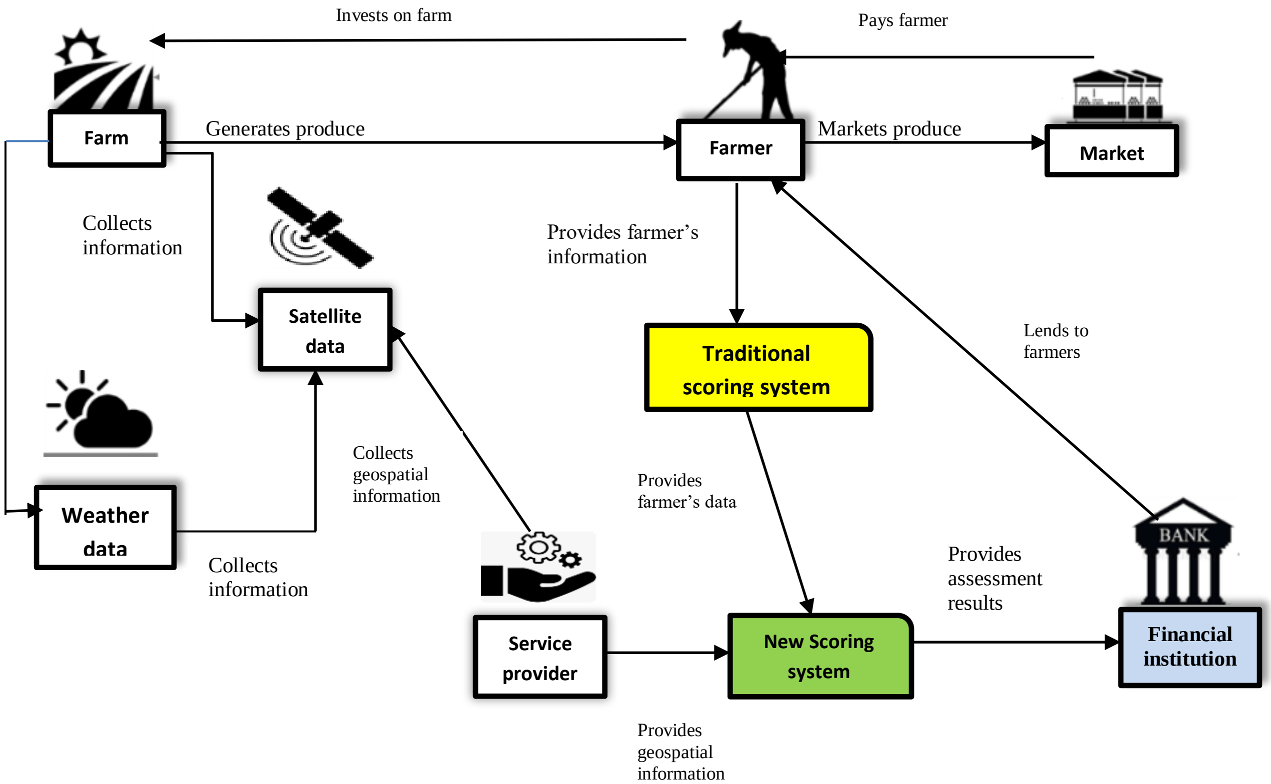

Figure 18: Proposed system institutional architecture

The setting up of this new service provider could start by looking at the present services at the Survey of Kenya (SOK), Department of Resource Surveys and Remote Sensing (DRSRS), Regional Center for the Mapping of Resources for Development (RCMRD), National Commission for Science, Technology and Innovation (NACOSTI) and related institutions.

5.2.3: Technical issues

The new scoring system would need to be able to receive whatever financial data there is for the farmers, together with crop geospatial data such as NDVI, according to the final approved model. It would then compute farmer credit scores using machine learning algorithms, and output them together with recommendations for credit access or denial, plus reasons. It would need a very friendly user interface for effective interaction with non-technical people. Figure 19 illustrates this proposed technical architecture.

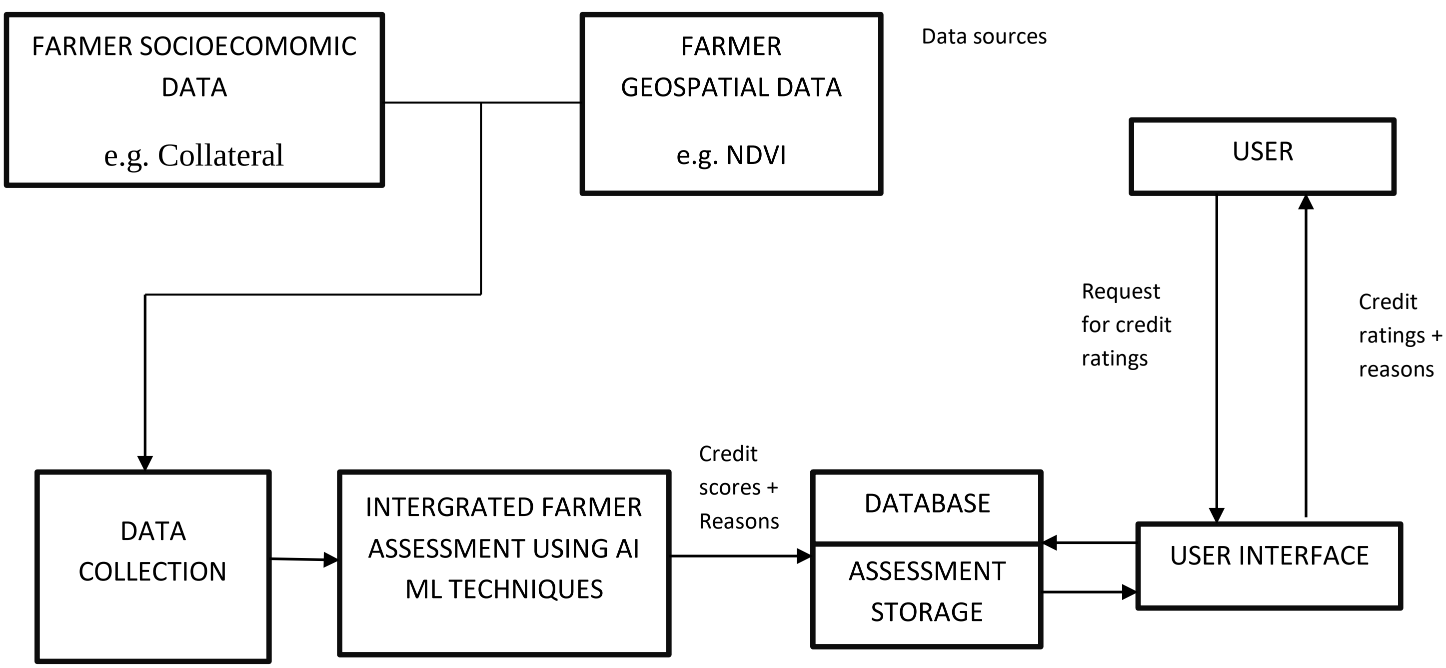

Figure 19: Proposed system technical architecture

5.2.4: Implementation steps

In the event that industry warms up to the idea presented in this research, the author envisages several systematic steps that could be followed towards implementation. These are summarized in Figure 20.

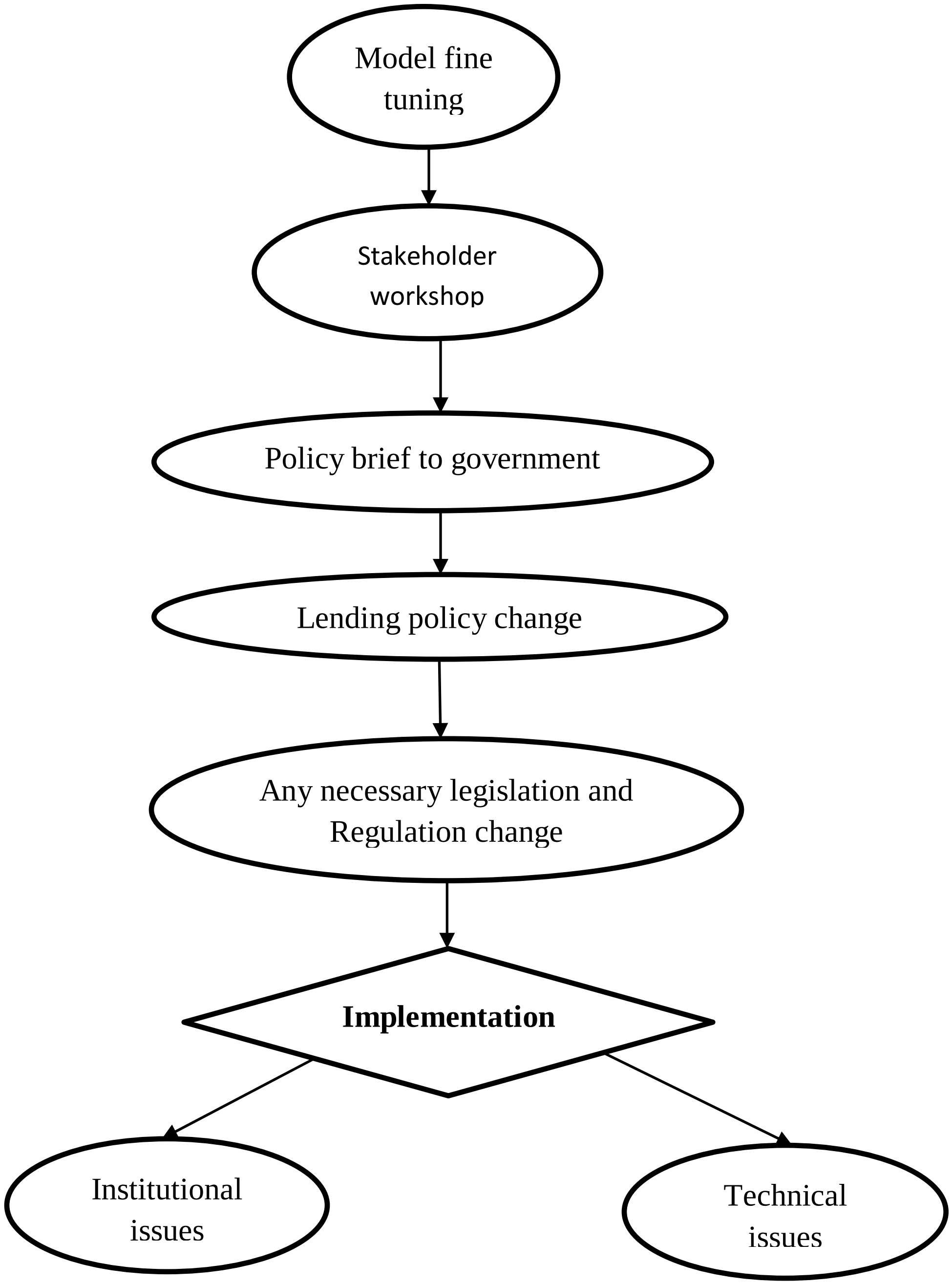

Figure 20: Possible implementation flowchart

6.0: CONCLUSIONS, RECOMMENDATIONS and CONTRIBUTION

6.1: Conclusions

The main objective of this study was to demonstrate how geospatial technology can be used to leverage farm credit scoring for the benefit of small holder farmers. In order to achieve this objective, a system was designed that would include geospatial data. This system would eventually improve on the chances of smallholder farmers' accessing loans to facilitate their farming activities. In order to achieve this, the following three specific objectives were tackled.

- To identify the small holder farms and farmers in the study area.
- To determine the extent of small holder farmer financial exclusion.
- To develop a new farmer credit scoring system that includes geospatial technology.

All these objectives have been achieved and the following conclusions are made:

- This research has established that indeed the majority of farmers in Migori County are smallholder farmers who produce the bulk of the staple maize that is consumed there.
- From the results achieved, the study demonstrated clearly that indeed, small holder farmers are financially excluded to a large extent in Migori County. As such, there is need to introduce new techniques for financially including such farmers in order to improve on their yields and hence on food security. Introduction of the use of non-financial data in lieu of collateral by financial institutions could assist these farmers towards financial inclusion.
- The research demonstrated that NDVI can be a useful tool in computing farmers' credit scores that can improve their chances of accessing loans from financial institutions.

6.2: Recommendations

From this study, it is recommended that:

- Financial institutions can adopt this new model following its fine tuning and wider testing; this is because the current system being used locks out many potential small holder farmers who may lack or not have sufficient collateral to offer, in the end affecting their scores and chances of accessing loans for their farming activities.
- Further studies can be conducted on how further geospatial variables (such as rainfall, evapotranspiration, etc.) could be used to improve and fine tune the model. Such further studies can also include further testing of the fine-tuned model in more geographic locations and agro climatic zones.

6.3: Contribution

The research findings from this study have contributed to the knowledge about credit scoring by demonstrating how geospatial technologies can be utilized to generate non-financial data as well as how such geo spatial data can complement historical information used by financial institutions to determine more farmer-friendly credit scores. The study has also contributed knowledge on the application of artificial intelligence machine learning techniques to agriculture, especially African agriculture. Finally, the study has contributed to the global debate on the financial inclusion of vulnerable small holder farmers whose farm outputs still remain critical to the achievement of food security in the world, and the realization of the UN sustainable development goal no.2 (no hunger).

7. REFERENCES

AFC Kenya, (2018). Credit Policy Manual 2018 Revision.

Albanis, G. T., and Batchelor R. A. (1999). Using probabilistic neural networks and rule induction techniques to predict long-term bond ratings, in M. Torres (Ed.), Proc. 5th Ann. Conf. Information Systems, Analysis and Synthesis, Orlando.

Alpaydin, E., (2010). Introduction to machine learning, 2nd edition. MIT Press, Cambridge, MA.

Alreck, P. L., and Settle, R.B. (1995). Survey Research Handbook. McGraw Hill, New York, 470 p. and its impact on Emerging Market Customers. Omidyar Network.

Barry, P. J. and Robinson, L. J. (2001). Agricultural Finance. In: Credit, Credit Constraints and Consequences. (Edited by Gardner, B. and Rausser, G.), Elsevier Ltd., Amsterdam.

Blockeel, H., De Raedt, L., and Ramon, J. (1998). Top-down induction of clustering trees. In: Proc. Fif - teenth International Conference on Machine Learning, p. 55–63. San Mateo, CA, Morgan Kaufmann

Cacoullos, T. (1966). Estimation of a multivariate density. Ann. Institute of Stat. Math. (Tokyo), 18: 179-189.

CEDA, (2006). Economic Empowerment for Botswana. Annual Report. Gaborone, Botswana.

Chavan, P., and Ramakumar, R. (2007). “Revival of Agricultural Credit in the 2000s:An Explaination”, *Economic and Political Weekly*, .Vol. XLII No.52

Chen, A., Leung M. T., and Daouk H. (2003). Application of neural networks for an emerging financial market: Forecasting and trading the Taiwan Stock Index. Comp. Operations Res., 30: 901-923.

Chigumira, G., and Masiyandima, N. (2003). Did Financial Sector Reform Result in Increased savings and lending for the SMEs and the poor.

Costa, A., Deb, A., and Kubzansky, M., (2016). Big Data, Small Credit: The Digital Revolution

Credit Reference Bureau Regulations, 2013. Kenya Gazette Supplement No. 3 (Legislative Supplement No. 3) (January 17, 2014)

Crook, J. N., Edelman D. E., and Thomas L. C. (2005). "Credit Scoring." Journal of the Operational Research Society 56, no. 9, pp. 1003-1005. Accessed April 04, 2017.

Etheridge, H. L., and Sriram R. S. (1997). A comparison of the relative costs of financial distress models: Artificial neural networks, logit and multivariate discriminant analysis. Intl. J. Intelligent Systems in Accounting, Finance and Management, 6: 235-248.

Farrin, Katie, and Mario J. Miranda, (2015). “A Heterogeneous Agent Model of Credit-Linked Index insurance and Farm Technology Adoption.” Journal of Development Economics 116, pp. 199–211.

Feder, G. (1987). Land registration and titling from an economist’s perspective: A case study in rural Thailand, survey review, vol. 29, No. 226, pp. 63-174.

Fielding, A. H. (1999). An introduction to machine learning methods. In: Fielding AH (ed) Machine Learning Methods for Ecological Applications, Kluwer Academic Publishers, Dordrecht.

FSD –Kenya Report (2013), Indexed Weather Insurance Initiatives.

Gestel, V., Tony and Baesens B. (2008). *Credit Risk Management: basic concepts: financial risk components, rating analysis, models, economic and regulatory capital*, Oxford University Press.

Government of Kenya (1969). The Agricultural Financial Corporation Act Chapter 323.

Gustafson, C. R., Bayer and Saxowsky, D. M. (1991). Credit Evaluation Investigating the Decision Process of Agricultural Loan Officers.” Agricultural Finance Review 51, pp. 55-63.

Gutiérrez-Nieto, Begoña, Carlos Serrano-Cinca, and Juan Camón-Cala, (2016). "A Credit Score System for Socially Responsible Lending." Journal of Business Ethics 133, no. 4, pp. 691- 701. Business Source Complete, EBSCOhost.

Hananu, B., Abdul-Hanan, A., Zakaria, H. (2015). Factors influencing agricultural credit demand in northern Ghana. African Journal of Agricultural Research 10, pp. 645–652.

Henley, E. (1995). Statistical aspects of credit scoring. Dissertation. The Open University, Milton Keynes, UK.

Henning, J. I. F., Jordaan, H. (2015). Investigating factors considered in agricultural credit applications, what are currently considered by commercial bank? In Proceedings of the 20th International Farm Management Congress, Quebec City, QC, Canada, 12–17, pp. 138–145.

Henning, J. I. F., Jordaan, H. (2016). Determinants of financial sustainability for farm credit applications, A Delphi Study. Sustainability 2016, vol 8, pp 77. https://www.farmcreditfunding.com/farmcredit

Icici applies satellite imaging to rural credit scoring; Article available at: https://www.finextra.com/newsarticle/36454/icici-applies-satellite-imaging-to-rural-credit-scoring

Jensen, L. (1992), Using neural networks for credit scoring. *Managerial Finance*; 18:15, 26.

Kaffenberger, M., and Chege, P., (2016). Digital Credit in Kenya: Time for Celebration or Concern? CGAP, 2016.

Kenya National Bureau of Statistics, (2019): The Kenya National Bureau of Statistics: Kenya Population and Housing Census.

Lacher, C., Coats, K., Sharma, S., and Fant, F. (1995). A neural network for classifying the "financial health of a firm. *European Journal of Operational Research,* 85:53, 65.

Lawal, J. O., Omonona, B. T., Ajani, O. I. Y., and Oni, O. A. (2009). Determinants of constraints to credit access among cocoa farming households in Osun State, Nigeria. Journal of Social Sciences 6.

Lee, T. S., Chiub, C. C., Luc, C. J., and Che, I..F. (2002). Credit scoring using the hybrid neural discriminant technique. Expert System Application, vol 23, pp. 245–254.

Limsombunchai, Christopher Gan and Minsoo Lee. (2005). American Journal of Applied Sciences 2 (8): pp.1198-1205.

Liverman. D., Moran, E. F., Rindfuss, R. and Stern, P.C. (Eds.). (1998). People and Pixels: Linking Remote Sensing and Social Science. National Academy Press, Washington DC.

Llanto. G. (2005). Rural Finance in the Philippines. Agriculture Credit Policy Council and Philippine. Institute for Development Studies, Philippine. 67.

McEvoy, M. (2013). "Enabling financial inclusion through 'alternative data.' MasterCard Advisors.

Mensah, S., (2004). A Review of SME Financing Schemes in Ghana, UNIDO Regional Workshop of Financing Small and Medium-scale Enterprises, pp. 3-20.

Mester, J. (1995). What's the point of credit scoring? *Business Review-Federal Reserve Bank of Philadelphia;* 3,16.

Ministry of Agriculture, (2006). Agricultural Support Schemes Guidelines. Agriculture Report, Gaborone, Botswana. September.

Monke, J. (2015). Agricultural Credit: Institutions and Issue, Congressional Research Services Report. April /8 /2015.

Moore, D., and McCabe, G. (1999). Introduction to the Practice of Statistics. W.H. Freeman and Company, New York, 823 p.

National Development Bank (NDB), (2006). Annual Report. Gaborone, Botswana.

Ngare, P., Kweyu, M., and Chris Huka, (2015). Modeling Risk of Financing Agribusiness in Kenya, Kenya Bankers Association, WPS/03/15.

Okeyo, S., Mulaku, G., and Mwange, C. (2022). Statistical Analysis of Small Holder Farmer Financial Exclusion: Case Study of Migori County, Kenya. *Open Journal of Statistics*, **12**, 733-742.

Ozili, P. K. (*2021*). "Financial inclusion and legal system quality: are they correlated?", Journal of Money and Business, Vol. 1 No. 2, pp. 84-101

Pal, D., and Sapre, A. (2010). “How have Government Policies Driven Rural Credit in India”, Indian Institutie of Management, Ahemedabad, W.P. No. 2010-12-02.

Parzen, E. (1962). An estimation of a probability density function and mode. Ann. Math. Stat., 33:1065-1076.

Peck, Robert, Chisten and Jamie Anderson. (2013). “Segmentation of Smallholder Households: Meeting the Range of Financial Needs in Agricultural Families.” CGAP. Focus Note No. 85.

Piramuthu, S. (1999). Financial credit-risk evaluation with neural and neurofuzzy systems. *European Journal of Operational Research*, 112:310, 21.

Quinlan, J. R. (1986). Induction of decision trees. Machine Learning 1:81-106.

Quainoo, T. K. (2011). The Impact of Loans on SMEs in Ghana, Masters Dissertation, Institute of Distance Learning, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, pp. 64-73.

Ramakumar R., and Chavan P. (*2007*). Revival in agricultural credit in the 2000s: An explanation. Economic and Political Weekly, 42(52), 57–64.

Rice, J. C. (1994). "Logistic regression: An introduction". In B. Thompson, ed., *Advances in social science methodology*, Vol. 3: 191-245. Greenwich, CT: JAI Press. Popular introduction

Rukuni, M. (2006). Revisiting Zimbabwe’s Agricultural Revolution. In Zimbabwe’s Agricultural Revolution Revisited, edited by Rukuni, M. Tawonezvi, P and Eicher, C.K. with Munyuki-Hungwe, M. and Matond, P. University of Zimbabwe Publications, Sable Press Private Limited, Harare, Zimbabwe, pp. 1-26, University of Zimbabwe, Harare, Zimbabwe.

Sacerdoti, E. (2005). Access to Bank Credit in Sub-Saharan Africa: Key Issues and Reform Strategies. IMF Working Paper, WP/05/166, Monetary and Financial Systems Department, August 2005.

Saima, A., and Hussain, Z. (2011). "Impact of Institutional Credit on Production Efficiency of Farming Sector", *Pakistan Economic and Social Review*, Volume 49, No.52, pp149-162

Salami, A., Kamara, A. B., and Brixiova, Z. (2010). Smallholder Agriculture in East Africa: Trends, Constraints and Opportunities. Working Papers Series No. 105, African Development Bank, Abidjan.

Satyasai, K. J. S. (2012). ”Equity in Agricultural Credit Delivery”, Paper presented at the 12th Annual Money and Finance Conference, Indira Gandhi Institute for Development Research, Mumbai.

Schleiter, I. M., Borchardt, D., Wagner, R., Dapper, T., Schmidt, K. D., Schmidt, H. H., and Werner, H. (1999). Modelling water quality, bioindication and population dynamics in lotic ecosystems using neural networks. Ecological Modelling 120: 271-286.

Schreiner, M. (2003). Scoring: the next breakthrough in microcredit? (English). CGAP Occasional paper; no. 7. Washington, DC: World Bank.

Shukla J. H., Jha P. J. (2018). Factors of bio-rhythms and applications to *PCA. Int J Stat Appl Math;3 (2):484-491.

Siles, M., Hanson, S. D., and Robison, L. J. (1994). Socio Economics and the Probability of Loan Approval Review of Agricultural Economics vol. 16 (2), pp. 363-372.

Simumba, N. (2018). Design and Evaluation of a Credit Decision System using Mobile Data in Agricultural Micro-finance, M.sc Dissertation, Keio University, Japan, 102 pp.

Specht, D.F. (1990). Probabilistic neural networks. Neural Networks, 3: 109-118.

The World Bank Group (World Bank) and GPFI, (2016). “Digital Financial Solution to Advance Women’s Economic Participation.” http://www.worldbank.org/en/programs/globalfindex

Thomas, L. C., David, B. E., and Jonathan, N. C. (2004). Readings in Credit Scoring: Foundations, Developments, and Aims. Oxford: Oxford University Press.

Toffler, A. (1985). The Adaptive Corporation, McGraw-Hill Inc, New York.

Udry, C. (2000). “Gender, Agricultural Production, and the Theory of the Household.” Readings in Development Microeconomics: 104.

Vens, C., Struyf, J., Schietgat, L., Džeroski, S., and Blockeel, H. (2008). Decision trees for hierarchical multi-label classification. Machine Learning 73:185-214.

Venugopal, V., and Baets W. (1994). Neural networks and statistical techniques in marketing research: A conceptual comparison. Marketing Intelligence & Planning, 12: 30-38.

Virendra, Pandi. (2018). Harvesting Inc to plough remote sensing data, AI-backed algorithms to generate credit scores for farmers, Published on March 28, 2018. Available on https://www.thehindubusinessline.com/companies/farm-loans-harvesting-inc-data-to-psbs-farmers-credit-scores/article23372588.ece

Weston, L. P. (2012). Your Credit Score: How to improve the 3-digit Number That Shapes Your Financial Future. 4th. Upper Saddle River, New Jersey: FT.

World Bank (2020) Digitization of Agribusiness Payments in Africa: Building a ramp for Farmers'financial inclusion and participation in a digital economy (http://documents.worldbank.org/en/915271601013162558/

Yang, Z. R., Platt M. B., and Platt H. D. (1999). Probabilistic neural networks in bankruptcy prediction. J. Business Res., 44: 67-74.

Zellar, M. and Sharma M. (1998). Rural Finance and Poverty Alleviation, No. 8; International Food Policy Research Institute (IFPRI): Washington, DC, USA.

Zumbika, N. (2000). The Challenges of Lending to Smallholder Farmers in Zimbabwe: A Personal Experience of Lending to Smallholder Farmers in Zimbabwe. Association Companies, Branches and Representatives throughout the World, Kaycee Printers Harare, Zimbabwe.

Websites:

https://www.myfico.com/

https://www.annualcreditreport.com/

https://www.listendata.com/

Appendix 1: Questionnaire

University of Nairobi

Dear Sir/Madam

I Susan Akello Okeyo, a Ph.D. student at the University of Nairobi, is conducting a survey on "**leveraging geospatial technologies for small holder farmer credit scoring**", as a part of my academic research. Please assist by completing this questionnaire.

The responses to this questionnaire will be used for academic purpose only and will assist in developing general findings and conclusions without specific reference to clients or institution, except where permission has been granted or where information may be available in the public domain.

PART A: FOR FARMERS

*Please check the box ☒

1. Personal information

Name: (optional):

Identification number:………………………….

Sex: Male ☐ Female ☐

Age of respondent: below 40: 40-49: ☐ 50-59: ☐ 60-69: ☐ 70+: ☐

Do you have a bank account? No ☐ Yes ☐

2. Marital status:

☐ Married

☐ Single

☐ Widow/Widower

☐ Divorced/ Separated

3. Occupation:

What type of farming activity do you conduct?

☐ Crop farming

☐ Livestock and crop farming

☐ Fish farming

☐ Crop, livestock and fish farming

4. Years of Occupation:

How long have you been farming?

1. 10-19 ☐ 2. 20-29 ☐ 3. 30-39 ☐ 4. 40-49 ☐ 5. 50+ ☐

5. How much land holding do you have.......... (1 Acre = 43560m^2)

6. Land ownership:

What ownership do you have on the land you are farming on?

☐ Personal land with the title deed under your name

☐ Family land

☐ Rented land

7. Where is your land located?

Location? ……………………………………..

Sub location? …………………………………

Village? ....

Land Reference number?..

8. Do you have more than one parcel of land? Yes ☐ No ☐

If yes how many parcels do you have?..................

9. Is your land freehold or leasehold? ....

10. Market

How and where do you sell your maize produce?

☐ Market

☐ Through brokers

☐ Direct from the farm

11. Credit

Have you ever applied for a loan to facilitate your farming activities?

Yes ☐ No ☐

If **yes**, what are the reasons?

☐ Fertilizers

☐ Seeds

☐ Pay workers

☐ Others (specify?)

If **no**, what are the reasons?

☐ No collateral

☐ Fear of losing collateral

☐ Others (specify?)

12. Collateral

What do financial institutions require as collateral before granting loans?

☐ Title deed

☐ Log book

☐ Crops in your farm

☐ Others (specify?)

13. Credit access

What are some of the challenges you faces when applying for loans?

☐ No collateral

☐ Low credit scores (rejected application)

☐ Lack of credit history

☐ Others (specify?)

14. Credit repayment

(a) If you have previously applied for a loan (loans), did you pay back on time?

Yes ☐ No ☐

If yes, how long did it take to pay back the loan/s?..

If no specify the reasons:

☐ Crop failure due to drought/floods

☐ Inability to pay back because you are fulfilling family needs

☐ Waiting for loan subsidies from government due to drought/floods

☐ Others (specify?)

(b) Where do you acquire you loan from?

☐ Bank

☐ Cooperatives

☐ Merry-go round (Chamas)

☐ Others

(c) What is your credit capacity? (Pay back capacity)

☐ 10,000-100,000 ☐ 100,000-250,000

☐ 250,000-500,000 ☐ 500,000-1,000,000

☐ 1,000,000 +

(d) For how long have you been taking loans over the years?...

(e) Do you have any outstanding loans? No ☐ Yes ☐

If yes, when did you apply for it?

When is it due?

How much have you offset so far?

15. Are you currently employed? No ☐ Yes ☐

If yes, please indicate your employer…………………………………………….

If no, please indicate whether you have a source/s of income?...

…………………………………………………………………………………………………

16. Please indicate the number of crop cycle per year in your farm…………………………..

17. What types of pest attacks do you encounter in your farming activities?..........................

………………………………………………………………………………………………

18. Do you use any organic pest control techniques?

No ☐ yes ☐

If yes please specify………………………………………………………………………

19. Do you use any type of irrigation techniques?

No ☐ yes ☐

If yes specify……………………………………………………………………………….

20. Which Ploughing methods do you use in farming?..

…………………………………………………………………………………………………

PART B: FOR FINANCIAL INSTITUTIONS

1. Personal information

Name: (optional): ……………………………………………………………………

Sex: Male ☐ Female ☐

Profession: ……………………………………………………………………………

Job title: ………………………………………………………………………………

2. Does your institution offer credits to smallholder farmers? Yes ☐ No ☐

If not, why? ....

…………………………………………………………………………………………

If yes, what criteria do you use to score and advance loans to farmers?

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

3. Does your institution use any scoring model to compute credit scores for farmers?

Yes ☐ No ☐

If yes, please indicate……………………………..

If no, what criteria do you use to score farmers?..

……………………………………………………………………………………………..

……………………………………………………………………………………………..

4. Do you use geospatial techniques (For example map, pointing, area) in developing credit scores for farmers?

Yes ☐ No ☐

If yes, please indicate which ones you use……………………………………………………

……

5. What are some of the challenges your institution faces when implementing credit scores for smallholder farmers?

☐ No credit history

☐ No collateral (e.g. title deeds)

☐ High rate of default

6. In case a farmer defaults paying, how do you recover the loan?

Thank you very much for completing this questionnaire! I appreciate your inputs; they will go a long way to help me finalize this project.

Appendix 2: Start variable correlation results

```
##      Serial Gender Age BankAcc Marital Occupation
## 1 20184113      2   1       1       3          1
## 2 27186633      2   1       2       1          1
## 3  8144389      1   3       2       1          2
## 4   221586      1   1       2       1          1
## 5 32261560      2   1       1       1          1
## 6 13388005      1   1       1       1          2
```

```
##            Serial Gender   Age BankAcc Marital Occupation Yearsofoccupatio
n
## Serial       1.00  -0.27  0.04    0.16   -0.09      -0.11             -0.0
6
## Gender      -0.27   1.00 -0.11   -0.11    0.17      -0.10              0.0
0
## Age          0.04  -0.11  1.00    0.17    0.13      -0.02              0.6
1
## BankAcc      0.16  -0.11  0.17    1.00   -0.04      -0.06              0.1
0
## Marital     -0.09   0.17  0.13   -0.04    1.00      -0.03              0.1
4
## Occupation  -0.11  -0.10 -0.02   -0.06   -0.03       1.00              0.0
1
##            Howmuchland Landownership Location Landparcels Howmany Landhold
## Serial            0.28         -0.06    -0.02       -0.15    0.30     0.00
## Gender           -0.13          0.00    -0.14        0.11   -0.14     0.12
## Age               0.18         -0.17    -0.13        0.03    0.04    -0.07
## BankAcc           0.09          0.12    -0.18       -0.02    0.04     0.04
## Marital          -0.07          0.00    -0.06        0.12   -0.15    -0.09
## Occupation        0.01          0.10     0.18       -0.04    0.05    -0.07
##            Market Credit Ifyes  Ifno Collateral Creditaccess Creditrepayme
nt
## Serial      -0.13  -0.05  0.14 -0.07      -0.04         0.10           -0.
09
## Gender      -0.05  -0.10  0.01 -0.18       0.01        -0.05            0.
00
## Age         -0.12  -0.06  0.12  0.14       0.11         0.03           -0.
04
## BankAcc     -0.08   0.00  0.15 -0.05       0.06         0.08           -0.
15
## Marital      0.01   0.01 -0.05  0.08      -0.02        -0.01           -0.
02
## Occupation   0.20   0.08 -0.14  0.21       0.04         0.05           -0.
09
##            Ifyess Reasons Acquire Capacity Howlong Outstanding Loanduratio
n
## Serial       0.00   -0.10   -0.08     0.24    0.14        0.12         0.0
```

```
3
## Gender       0.13   -0.06    0.11    -0.08   -0.03        0.03         0.0
2
## Age          0.06   -0.12   -0.03     0.09    0.18       -0.01        -0.0
2
## BankAcc      0.03   -0.16   -0.05     0.14    0.14        0.01         0.0
8
## Marital     -0.02    0.02   -0.05    -0.07   -0.04       -0.01         0.0
6
## Occupation  -0.03   -0.02   -0.04    -0.05   -0.07       -0.14        -0.0
8
##              Set Employment Employer Income Cycle  Pest Organic Specify
## Serial     -0.04       0.23     0.42  -0.05  0.11  0.19   -0.08   -0.11
## Gender      0.01      -0.15    -0.15  -0.09 -0.06 -0.03   -0.02    0.10
## Age        -0.04       0.02    -0.04   0.06  0.06 -0.07    0.03   -0.04
## BankAcc     0.03       0.20     0.16   0.07  0.15 -0.10    0.01    0.00
## Marital     0.02       0.02    -0.09  -0.02 -0.05 -0.08   -0.01    0.13
## Occupation -0.08       0.04    -0.01   0.09 -0.05 -0.02   -0.07   -0.01
##            Irrigation Specifyirrigation Ploughing credit perfomance  perf
## Serial          -0.04             -0.09      0.49     NA         NA  0.11
## Gender           0.04              0.11     -0.18     NA         NA  0.05
## Age              0.05              0.05      0.08     NA         NA  0.14
## BankAcc          0.10             -0.01      0.13     NA         NA  0.13
## Marital          0.10              0.13     -0.01     NA         NA -0.10
## Occupation      -0.16             -0.03     -0.02     NA         NA -0.03

##         Var1   Var2 value
## 1     Serial Serial  1.00
## 2     Gender Serial -0.27
## 3        Age Serial  0.04
## 4    BankAcc Serial  0.16
## 5    Marital Serial -0.09
## 6 Occupation Serial -0.11
```

Appendix 3: List of small holder farms and farmers

SERIAL NO	FARM ID (θ)	FARM LAT (λ)	FARM LONG	FARMER ID	FARM NDVI AVERAGE	SUB-COUNTY
1	17	-0.772	34.625	112548	0.58	Suna East
2	4	-0.88	34.543	221586	0.55	Rongo
3	102	-0.724	34.628	245812	0.60	Suna East
4	10	-0.937	34.521	362547	0.33	Suna West
5	42	-1.084	34.499	1013025	0.56	Rongo
6	60	-0.988	34.502	1076489	0.54	Rongo
7	68	-0.945	34.518	1121536	0.56	Rongo
8	216	-1.258	34.648	1847119	0.50	Kuria West
9	43	-1.088	34.501	2325116	0.58	Rongo
10	50	-1.084	34.587	2403160	0.53	Rongo
11	59	-0.962	34.387	2492236	0.50	Awendo
12	133	-1.088	34.447	29752960	0.35	Suna West
13	58	-0.967	34.384	2704242	0.50	Rongo
14	140	-1.024	34.324	2726065	0.46	Suna West
15	45	-1.095	34.523	2739590	0.57	Rongo
16	41	-1.09	34.414	2771059	0.55	Rongo
17	47	-1.1	34.546	2779794	0.61	Rongo
18	57	-0.965	34.37	2799334	0.49	Rongo
19	7	-0.84	34.563	2803912	0.54	Suna East
20	146	-1.07	34.332	2954713	0.45	Awendo
21	141	-1.022	34.29	2971650	0.47	Suna West
22	46	-1.095	34.534	3140933	0.59	Rongo
23	145	-1.057	34.318	3207642	0.48	Awendo
24	6	-0.850	34.563	13388005	0.74	Rongo
25	105	-0.708	34.563	3265412	0.52	Suna East
26	19	-0.758	34.626	3369874	0.61	Suna East
27	52	-1.048	34.512	3372444	0.59	Rongo
28	39	-1.089	34.355	3654821	0.45	Rongo
29	44	-1.091	34.509	3654954	0.58	Rongo
30	54	-0.995	34.404	3659453	0.52	Rongo
31	130	-1.048	34.452	3972355	0.53	Suna West
32	139	-1.032	34.346	3973488	0.48	Suna West
33	142	-1.019	34.285	3973824	0.43	Suna West

34	144	-1.029	34.306	3981961	0.43	Suna West
35	138	-1.038	34.358	4130255	0.51	Suna West
36	55	-0.991	34.402	4568952	0.53	Rongo
37	56	-0.989	34.397	5643871	0.53	Awendo
38	48	-1.088	34.563	6533713	0.60	Rongo
39	128	-1.155	34.447	6912045	0.55	Suna West
40	9	-0.935	34.522	6912420	0.33	Suna East
41	209	-1.351	34.707	23348738	0.45	Kuria West
42	49	-1.087	34.58	8126984	0.55	Rongo
43	3	-0.883	34.54	8144389	0.55	Rongo
44	40	-1.104	34.39	8159461	0.50	Rongo
45	218	-1.224	34.655	8953241	0.51	Kuria West
46	51	-1.052	34.538	9771763	0.57	Rongo
47	67	-0.913	34.54	10223359	0.52	Rongo
48	20	-0.728	34.629	10233261	0.60	Awendo
49	65	-0.955	34.558	10543276	0.55	Rongo
50	126	-1.177	34.432	11215846	0.54	Suna West
51	204	-1.313	34.621	11224046	0.54	Kuria West
52	221	-1.144	34.584	11412683	0.54	Kuria West
53	12	-0.889	34.595	13112933	0.53	Uriri
54	222	-1.139	34.57	13244463	0.52	Suna East
55	106	-0.711	34.569	13385066	0.56	Uriri
56	15	-0.817	34.591	13386264	0.56	Uriri
57	208	-1.347	34.704	14620284	0.55	Kuria East
58	72	-0.906	34.466	14677758	0.53	Rongo
59	1	-0.902	34.524	20184113	0.52	Rongo
60	125	-1.194	34.407	20658227	0.52	Suna West
61	16	-0.786	34.598	21086953	0.56	Suna East
62	14	-0.829	34.586	21186481	0.53	Uriri
63	64	-0.966	34.542	21549238	0.57	Awendo
64	13	-0.903	34.608	22211547	0.56	Suna East
65	70	-0.941	34.497	22304385	0.54	Rongo
66	69	-0.948	34.508	22407658	0.56	Rongo
67	129	-1.044	34.439	22724361	0.52	Suna West
68	209	-1.351	34.707	23348738	0.54	Kuria West
69	62	-1.006	34.551	23890276	0.54	Rongo
70	206	-1.331	34.645	23969259	0.52	Kuria East
71	18	-0.774	34.625	24244325	0.58	Suna East

72	71	-0.908	34.477	24503941	0.54	Rongo
73	61	-0.987	34.546	25274494	0.51	Rongo
74	217	-1.239	34.652	25696211	0.53	Kuria West
75	203	-1.298	34.615	26451306	0.54	Kuria East
76	135	-1.067	34.42	26811913	0.53	Suna West
77	2	-0.886	34.54	27186633	0.57	Rongo
78	136	-1.054	34.404	28304082	0.53	Suna West
79	205	-1.323	34.627	28563019	0.54	Kuria East
80	124	-1.224	34.479	28894120	0.52	Suna West
81	134	-1.078	34.443	29752960	0.56	Suna West
82	127	-1.175	34.433	30002630	0.55	Suna West
83	11	-0.882	34.591	30142286	0.52	Suna East
84	73	-0.903	34.477	30422183	0.54	Rongo
85	63	-1.017	34.564	31405366	0.55	Rongo
86	131	-1.052	34.463	31509272	0.53	Suna West
87	202	-1.269	34.612	31617838	0.51	Kuria East
88	219	-1.215	34.653	31702923	0.53	Kuria west
89	66	-0.944	34.557	32145128	0.48	Rongo
90	137	-1.04	34.38	32241069	0.52	Suna West
91	5	-0.857	34.566	32261560	0.53	Rongo
92	132	-1.092	34.457	32970842	0.52	Suna West
93	207	-1.334	34.676	33578430	0.53	Kuria East
94	75	-0.899	34.514	34296712	0.46	Rongo
95	104	-0.704	34.555	34385725	0.51	Suna East
96	220	-1.16	34.62	34599488	0.56	Kuria West
97	201	-1.263	34.613	35062878	0.53	Kuria West
98	8	-0.928	34.523	37703733	0.41	Suna East
99	74	-0.9	34.514	37752062	0.46	Rongo
100	103	-0.722	34.628	71285331	0.56	Suna East
101	210	-1.361	34.714	72139428	0.55	Kuria West

Appendix 4: Model training results

Farm serial	Farm id	Farmer ID	Land parcels	Collateral	Reasons	Howlong	NDVI	Previous Credit	Traditional score	With 1.0 NDVI	With 1.2 NDVI	With 1.3 NDVI	With 1.4 NDVI	With 1.5 NDVI
1	1	20184113	2	1	Null	Null	0.64	0	685	722	806	848	850	850
2	2	27186633	2	5	Null	Null	0.62	0	599	641	709	743	777	811
3	3	8144389	1	1	Null	Null	0.63	0	639	663	735	771	807	844
4	4	221586	2	1	Null	Null	0.62	0	685	722	806	848	850	850
5	5	32261560	2	Null	Null	Null	0.56	0	726	661	733	769	805	841
6	7	2803912	2	4	Null	20	0.52	1	367	408	430	441	452	463
7	8	37703733	2	2	Null	1	0.30	0	510	499	538	558	578	598
8	10	362547	2	Null	Null	Null	0.60	0	726	661	733	769	805	841
9	11	30142286	2	Null	Null	5	0.33	0	565	549	598	623	648	673
10	13	22211547	2	Null	Null	Null	0.58	0	726	661	733	769	805	841
11	15	13386264	1	4	1	10	0.65	0	462	474	509	526	543	561
12	17	112548	2	1	Null	2	0.58	0	521	447	476	491	506	520
13	20	10233261	2	7	Null	Null	0.61	0	637	676	751	789	827	850
14	41	2771059	1	Null	Null	Null	0.49	0	681	717	800	842	850	850
15	43	23251161	1	2	Null	5	0.68	0	449	463	495	512	528	544
16	44	365495	2	1	2	30	0.59	1	594	515	558	580	601	623
17	45	2739590	2	7	Null	Null	0.57	1	637	577	632	660	687	715
18	46	31409333	2	7	2	Null	0.46	0	777	833	850	850	850	850
19	46	31409333	2	7	2	Null	0.53	0	777	718	801	843	850	850
20	47	27797946	2	Null	2	1	0.66	0	720	745	834	850	850	850
21	48	6533713	1	Null	Null	Null	0.52	0	681	717	800	842	850	850
22	49	8126984	1	2	Null	10	0.56	1	379	300	300	300	300	300
23	54	365945	2	Null	Null	Null	0.59	0	726	661	733	769	805	841
24	55	4568952	1	5	Null	10	0.53	1	322	300	300	300	300	300
25	58	2774805	2	1	Null	Null	0.43	0	685	682	759	797	835	850
26	59	27042422	2	Null	Null	20	0.49	1	495	528	574	597	620	642

27	61	1076489	2	1	1	5	0.47	0	664	707	788	828	850	850
28	62	25274494	1	8	Null	9	0.24	1	430	405	426	437	447	458
29	64	31405366	2	1	1	6	0.60	0	664	691	769	808	847	850
30	65	21549238	2	1	3	1	0.68	0	679	707	788	828	850	850
31	66	10543276	2	1	2	10	0.67	0	594	615	678	709	740	772
32	67	32145128	2	Null	Null	Null	0.54	0	726	661	733	769	805	841
33	68	10223359	1	1	Null	4	0.56	1	428	336	344	347	351	355
34	69	1121536	2	1	2	30	0.52	0	594	631	697	730	763	796
35	73	14677758	1	1	Null	Null	0.51	0	639	678	754	792	830	850
36	74	30422183	2	1	Null	Null	0.48	0	685	738	825	850	850	850
37	102	245812	2	10	Null	10	0.63	1	406	428	454	467	479	492
38	103	7.13E+08	2	6	Null	2	0.46	1	435	425	451	463	476	488
39	104	34385725	2	10	Null	6	0.57	1	476	405	425	436	446	457
40	105	3265412	2	6	Null	Null	0.58	1	599	541	589	613	637	661
41	106	13385066	2	Null	Null	Null	0.55	0	726	661	733	769	805	841
42	124	20658227	2	3	Null	4	0.54	1	388	314	317	319	320	322
43	128	22724361	2	1	Null	5	0.42	0	524	510	552	573	594	615
44	130	31509272	2	1	Null	5	0.58	0	524	450	480	495	510	525
45	131	32970842	2	1	Null	3	0.48	0	474	511	553	575	596	617
46	133	29752960	2	Null	Null	Null	0.35	0	726	721	805	847	850	850
47	135	28304082	2	2	Null	7	0.58	1	495	423	447	460	472	484
48	137	4130255	2	8	Null	5	0.39	1	476	465	497	514	530	547
49	138	3973488	2	1	Null	3	0.68	1	474	495	534	554	573	593
50	139	27260651	2	1	Null	Null	0.50	0	685	738	825	850	850	850
51	140	29716500	2	Null	Null	Null	0.46	0	726	721	805	847	850	850
52	141	3973824	1	2	Null	5	0.44	1	449	423	448	460	473	485
53	142	28901345	1	Null	Null	Null	0.48	0	681	717	800	842	850	850
54	144	32076422	2	Null	Null	Null	0.43	0	726	721	805	847	850	850
55	146	7.21E+08	1	Null	Null	Null	0.37	0	681	661	734	770	806	842
56	201	35062878	2	Null	Null	Null	0.56	0	726	661	733	769	805	841
57	202	31617838	2	1	Null	Null	0.44	0	685	682	759	797	835	850

58	203	26451306	2	7	Null	3	0.53	1	426	350	360	365	370	375
59	205	28563019	2	1	Null	Null	0.59	0	685	622	687	719	751	783
60	206	23969259	2	1	Null	Null	0.46	0	685	738	825	850	850	850
61	207	33578430	2	1	Null	Null	0.42	0	685	682	759	797	835	850
62	208	14620284	1	5	Null	Null	0.32	0	553	542	590	614	638	662
63	209	23348738	1	1	Null	1	0.45	0	493	467	500	517	533	550
64	216	1847119	2	7	2	Null	0.36	0	777	777	850	850	850	850
65	218	8953241	2	1	2	Null	0.41	0	825	823	850	850	850	850
66	219	31702923	2	2	2	Null	0.61	0	796	835	850	850	850	850
67	220	34599488	Null	1	2	1	0.34	1	652	632	699	732	765	798

Appendix 5: Descriptive Analysis - Questionnaires

The analysis showed that out 400 respondents interviewed in Migori County, 57% were male and 43% were female.

Gender Analysis

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	Male	181	56.7	56.9	56.9
	Female	137	42.9	43.1	100.0
	Total	318	99.7	100.0	
Missing	System	1	.3		
Total		**319**	**100**		

Age stratification:
51% of the total respondents were below the age of 40 years, 18.8% were between the ages of 50 to 59 years, 12.5% were between ages of 60 to 69 years and 2.2% were above 70 years of age.

Age of respondents

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	Below 40	163	51.1	54.7	54.7
	40-49	60	18.8	20.1	74.8
	50-59	40	12.5	13.4	88.3
	60-69	28	8.8	9.4	97.7
	70+	7	2.2	2.3	100.0
	Total	298	93.4	100.0	
Missing	System	21	6.6		
Total		319	100.0		

Ownership of Bank Accounts
The analysis further revealed that 57.4% of the total respondents did not have bank accounts, while 32% of the respondents had bank accounts.

Do you have bank account

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	No	183	57.4	64.2	64.2

	Yes	102	32.0	35.8	100.0
	Total	285	89.3	100.0	
Missing	System	34	10.7		
Total		319	100.0		

82% of the respondents were married, 9.1% were

Marital Status

Results indicated that of the 400 respondents interviewed, 81.5 % were married, 9.1% were single, 8.5% were widowed and 0.3% are divorced/separated.

Marital status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Married	260	81.5	82.0	82.0
	Single	29	9.1	9.1	91.2
	Widow/Widower	27	8.5	8.5	99.7
	Divorced/Separated	1	.3	.3	100.0
	Total	317	99.4	100.0	
Missing	System	2	.6		
Total		319	100.0		

Farming Activity:

53% of respondents practice crop farming, 41.3% practice both crop and livestock farming, 0.3% do fishing and 5.4 practice crop, livestock and fish farming.

What type of farming activity do you conduct?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Crop farming	168	52.7	53.0	53.0
	Livestock and crop farming	131	41.1	41.3	94.3
	Fish farming	1	.3	.3	94.6
	Crop, livestock and fish farming	17	5.3	5.4	100.0
	Total	317	99.4	100.0	
Missing	System	2	.6		

Total	319	100.0		

Duration of Practicing the Occupation

57.6% respondents had been practicing farming for a period between 10 to 19 years, 26.1% have been practicing farming between 20 to 29 years, 8.3% have been farming between 30 to 39 years, 6.7% have been practicing farming between 40 to 49 years while 1.3% had been practicing farming for more than 50 years. The analysis showed that most farmers have been farming for less than 20 years.

Duration of Years of practicing the occupation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10-19 years	181	56.7	57.6	57.6
	20-29 years	82	25.7	26.1	83.8
	30-39 years	26	8.2	8.3	92.0
	40-49 years	21	6.6	6.7	98.7
	50+	4	1.3	1.3	100.0
	Total	314	98.4	100.0	
Missing	System	5	1.6		
Total		319	100.0		

Land Ownership:

The minimum land holding is 0.2 acres while the maximum land holding is 60 acres. The mean of all land holdings is 3.1 with a standard deviation of 6.8.

Standard deviation shows how the individual scores deviate from mean.

How much land holding do you have?

	N	Minimum	Maximum	Mean	Std. Deviation
how much land holding do you have	263	.20	60.00	3.1521	6.76436
Valid N (listwise)	263				

64.2% of respondents farm on family land. 20.3% farm on their personal land with the title deed in their names, 13% farm on leased land and 2.2% farm on both family and rented land.

Land ownership

	Frequency	Percent	Valid Percent	Cumulative Percent

Valid	Personal land with the title deed under your name	64	20.1	20.3	20.3
	Family land	203	63.6	64.2	84.5
	Rented land	42	13.2	13.3	97.8
	Family and rented land	7	2.2	2.2	100.0
	Total	316	99.1	100.0	
Missing	System	3	.9		
Total		319	100.0		

Location of farm in Migori County

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Suna lower	26	8.2	8.5	8.5
	Nyabisawa	11	3.4	3.6	12.1
	Otacho	8	2.5	2.6	14.7
	East Kamagambo	6	1.9	2.0	16.6
	Central Kamagambo	17	5.3	5.5	22.1
	North Kamagambo	6	1.9	2.0	24.1
	West Kamagambo	10	3.1	3.3	27.4
	South Kamagambo	17	5.3	5.5	32.9
	South Sakwa	15	4.7	4.9	37.8
	Kamagambo	6	1.9	2.0	39.7
	Suna South	17	5.3	5.5	45.3
	Anjego	1	.3	.3	45.6
	Kakrao	1	.3	.3	45.9
	Osingo	9	2.8	2.9	48.9
	Nyamongo	1	.3	.3	49.2
	God jope	3	.9	1.0	50.2
	Kwa	1	.3	.3	50.5
	Suna west	3	.9	1.0	51.5
	Central Kanyamkago	4	1.3	1.3	52.8
	East Sakwa	11	3.4	3.6	56.4
	Central Sakwa	6	1.9	2.0	58.3

North Sakwa	1	.3	.3	58.6
Alego west	1	.3	.3	59.0
South Kanyamkago	15	4.7	4.9	63.8
South East Kanyamkago	1	.3	.3	64.2
Kanyamkago	5	1.6	1.6	65.8
Tebesi	1	.3	.3	66.1
Nyabikongori	1	.3	.3	66.4
Nyaroha	2	.6	.7	67.1
Nyabasi South	3	.9	1.0	68.1
Komotobo	3	.9	1.0	69.1
Kebaroti	1	.3	.3	69.4
Nguruna	1	.3	.3	69.7
Kemakoba	1	.3	.3	70.0
Mabera	3	.9	1.0	71.0
Masurura	4	1.3	1.3	72.3
Nyamosense	6	1.9	2.0	74.3
Kumumwamu	2	.6	.7	74.9
Bugumbe South	4	1.3	1.3	76.2
Bugumbe West	2	.6	.7	76.9
Komosoko	3	.9	1.0	77.9
Kengarisio	1	.3	.3	78.2
Makerero	9	2.8	2.9	81.1
Gwitembe	3	.9	1.0	82.1
Bwirege Central	4	1.3	1.3	83.4
Bwirege East	3	.9	1.0	84.4
Siabai	1	.3	.3	84.7
Ikegere	18	5.6	5.9	90.6
Bukira west	12	3.8	3.9	94.5
Nyabikaye	4	1.3	1.3	95.8
Bukira	1	.3	.3	96.1
Mashangwe	1	.3	.3	96.4
Kwihu	1	.3	.3	96.7
Kilimakebe	2	.6	.7	97.4
Taragwiti	1	.3	.3	97.7
Ntimaru	1	.3	.3	98.0
Mashangwe	1	.3	.3	98.4

	Kugitimo	2	.6	.7	99.0
	Maeta	1	.3	.3	99.3
	Chinato	1	.3	.3	99.7
	Getabwega	1	.3	.3	100.0
	Total	307	96.2	100.0	
Missing	System	12	3.8		
Total		319	100.0		

76.4% of respondents said that they only one had one parcel of land while 23.6% have more than one parcel of land.

Do you have more than one parcel of land

		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Yes	73	22.9	23.6	23.6
	No	236	74.0	76.4	100.0
	Total	309	96.9	100.0	
Missing	System	10	3.1		
Total		319	100.0		

For those respondents who have more than one parcel of land, 37.5% have 1 more parcel of land, 36.5% have two more parcels, 15.6 % have three more parcels, 5.2% have 4 and 5 more parcels.

If yes how many parcel

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	36	11.3	37.5	37.5
	2	35	11.0	36.5	74.0
	3	15	4.7	15.6	89.6
	4	5	1.6	5.2	94.8
	5	5	1.6	5.2	100.0
	Total	96	30.1	100.0	
Missing	System	223	69.9		
Total		319	100.0		

80% of farmers have freehold lands while 19.1% have leasehold lands.

Is your land freehold or leasehold?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Freehold	246	77.1	80.9	80.9
	Leasehold	58	18.2	19.1	100.0
	Total	304	95.3	100.0	
Missing	System	15	4.7		
Total		319	100.0		

60% of farmers sell their produce direct to the market, 27.7% sell through brokers, 8.3% sell direct from their farms and 3.8% sell through brokers and to the market.

Market for Selling Maize Produce

How and where do you sell your maize produce

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Market	189	59.2	60.2	60.2
	Through brokers	87	27.3	27.7	87.9
	Direct from the farm	26	8.2	8.3	96.2
	Market and Brokers	12	3.8	3.8	100.0
	Total	314	98.4	100.0	
Missing	System	5	1.6		
Total		319	100.0		

Loan Facilities:

The analysis show that 35% of the farmers had never applied for loans while 65% have.

Have you ever applied for a loan to facilitate your farming activities?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	99	31.0	35.0	35.0
	No	184	57.7	65.0	100.0
	Total	283	88.7	100.0	
Missing	System	36	11.3		
Total		319	100.0		

51.4% who apply for loans do so to buy fertilizers and seeds,19.9% use loans to buy fertilizers, 11.9% use loans to buy seeds, 1.8% use loans to pay workers, 0.9% use loans to buy pesticides

This shows that most farmers apply for loans to buy fertilizers and seeds according to the analysis.

If yes what are the reasons

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Fertilizers	21	6.6	19.3	19.3
	Seeds	13	4.1	11.9	31.2
	Pay workers	2	.6	1.8	33.0
	Fertilizer and to buy pesticide	1	.3	.9	33.9
	Fertilizer and seeds	56	17.6	51.4	85.3
	Fertilizer, seeds and pay workers	14	4.4	12.8	98.2
	Business	2	.6	1.8	100.0
	Total	109	34.2	100.0	
Missing	System	210	65.8		
Total		319	100.0		

44.8% of respondent who do not apply for loans do not have collateral to offer to the financial institutions. 36.6% have collateral but fear losing them. 7.2% don't apply because of lack of awareness, 8.8% don't apply because they are financially stable and 0.5% don't apply because they lack guarantors

If no what are the reasons?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No collateral	87	27.3	44.8	44.8
	Fear of losing collateral	71	22.3	36.6	81.4
	Lack of awareness	14	4.4	7.2	88.7
	Personal	1	.3	.5	89.2
	Due to process	2	.6	1.0	90.2
	Guarantors	1	.3	.5	90.7
	Financially stable	17	5.3	8.8	99.5
	No collateral and fear of losing collateral	1	.3	.5	100.0
	Total	194	60.8	100.0	
Missing	System	125	39.2		
Total		319	100.0		

58.1% of respondents indicated that title deed were required as collateral by financial institutions, 9.8% indicated that log books were required, 3.4% indicated that crops were required as collateral, 15.8% indicated title deed and log books and 2.1% indicated guarantors and deposit.

What do financial institutions require as collateral before granting loans?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Title deed	136	42.6	58.1	58.1
	Log book	23	7.2	9.8	67.9
	Crops in your farm	8	2.5	3.4	71.4
	Bank Account	3	.9	1.3	72.6
	Title deed and Crops in your farm	5	1.6	2.1	74.8
	Deposit	6	1.9	2.6	77.4
	Title and log book	37	11.6	15.8	93.2
	Guarantor and deposits	5	1.6	2.1	95.3
	ID number and deposit	6	1.9	2.6	97.9
	Property	5	1.6	2.1	100.0
	Total	234	73.4	100.0	
Missing	System	85	26.6		
Total		319	100.0		

32.7% of respondents face the challenge of not having any collateral when applying for loans. 26.1% have low credit scores hence rejected during the process of application. 20.4% lack credit history, 0.9% lack guarantors and 0.4% lack registration fee.

What are some of the challenges you face when applying for loans

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No collateral	74	23.2	32.7	32.7
	Low credit scores (rejected application)	59	18.5	26.1	58.8
	Lack of credit history	46	14.4	20.4	79.2
	Lack of interest	1	.3	.4	79.6

	Lack of collateral and registration fee	1	.3	.4	80.1
	Guarantor	2	.6	.9	81.0
	No collateral and low credit scores	31	9.7	13.7	94.7
	Lack of registration fee	1	.3	.4	95.1
	Short repayment period	2	.6	.9	96.0
	Defaulting	1	.3	.4	96.5
	Delay	4	1.3	1.8	98.2
	No collateral and lack of credit history	4	1.3	1.8	100.0
	Total	226	70.8	100.0	
Missing	System	93	29.2		
Total		319	100.0		

50.5% of respondents who have previously applied for loans indicated that they had not paid back while 49.5% had paid back their loans.

If you have previously applied for a loan, did you repay back?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	95	29.8	49.5	49.0
	No	97	30.4	50.5	99.5
					100.0
	Total	192	60.2	100.0	
Missing	System	127	39.8		
Total		319	100.0		

The minimum time it took for some respondents to pay back their loans was one month while the maximum time they took was 24 months. The mean was 4.28, meaning that it took most respondents a period of 4 months to pay back their loans.

If yes how long did it take you to pay back?

	N	Minimum	Maximum	Mean	Std. Deviation
If yes how long did it take to pay back the loans	93	1	24	4.28	3.639

Valid N (listwise)	93					

For those respondents who had not paid back their loans, 65.1% indicated that this was due to the fact that they were fulfilling their family needs first, 22.1% indicated that it was due to drought/floods leading to crop failure, 7% were waiting for loan subsidies from government due to draught/floods and 5.8% was due to Crop failure and inability to pay back.

If no specify

		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Crops failure due to drought/floods	19	6.0	22.1	22.1
	Inability to pay back because you are fulfilling family needs	56	17.6	65.1	87.2
	Waiting for loan subsidies from government due to draught/floods	6	1.9	7.0	94.2
	Crop failure and inability to pay back	5	1.6	5.8	100.0
	Total	86	27.0	100.0	
Missing	System	233	73.0		
Total		319	100.0		

Source of Loan

42.7% of the respondents acquire loans from banks, 21.4% acquire from cooperatives, 29,7% acquire loans from village saving groups (*chamas)* and 2.6% acquire loans from an organization called One Acre Fund.

Where do you acquire your loan from?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Bank	82	25.7	42.7	42.7
	Cooperative	41	12.9	21.4	64.1
	Merry-go round (Chamas)	57	17.9	29.7	93.8
	Personal support	2	.6	1.0	94.8

	One Acre fund	5	1.6	2.6	97.4
	Merry go round and one acre fund	1	.3	.5	97.9
	Digi farm	1	.3	.5	98.4
	Nuru	1	.3	.5	99.0
	Uwezo	1	.3	.5	99.5
	Community based organizations	1	.3	.5	100.0
	Total	192	60.2	100.0	
Missing	System	127	39.8		
Total		319	100.0		

Credit Capacity

86.7% 0f the total respondents' credit capacity lay at a range of between 10,000 and 100,000, 7.9% between 100,000 and 250,000, 3.7% between 250,000 and 500,000 and 0.5% at over 1, 000,000.

What is your credit capacity (Pay back capacity)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10000-100000	165	51.7	86.4	86.4
	100000-250000	15	4.7	7.9	94.2
	250000-500000	7	2.2	3.7	97.9
	500000-1000000	1	.3	.5	98.4
	1000000+	1	.3	.5	99.0
	Less than 10000	2	.6	1.0	100.0
	Total	191	59.9	100.0	
Missing	System	128	40.1		
Total		319	100.0		

92.8% of the respondents indicated that they did not have any outstanding loans, while 7.2% had outstanding loans.

Do you have any outstanding loans?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	218	68.3	92.8	92.8
	Yes	17	5.3	7.2	100.0

	Total	235	73.7	100.0	
Missing	System	84	26.3		
Total		319	100.0		

Frequency of Applying for Loans
38.5% of respondents had taken loans during the past one year, 15.4% had taken loan in the past 10 months, 7.7% had taken loans in the past 5 months and 7.7% had taken loans in the past 2 years.

Loan duration

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 year	5	1.6	38.5	38.5
	10 months	2	.6	15.4	53.8
	5 months	1	.3	7.7	61.5
	1 year 6 months	1	.3	7.7	69.2
	4 years	1	.3	7.7	76.9
	2 years	1	.3	7.7	84.6
	1 month	1	.3	7.7	92.3
	8 months	1	.3	7.7	100.0
	Total	13	4.1	100.0	
Missing	System	306	95.9		
Total		319	100.0		

Status of employment:
89.5% of respondents are currently employed. 9.9% are not employed and 0.6% have retired.

Are you currently employed?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	281	88.1	89.5	89.5
	No	31	9.7	9.9	99.4
	Retired	2	.6	.6	100.0
	Total	314	98.4	100.0	
Missing	System	5	1.6		
Total		319	100.0		

If yes please indicate your employer

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	TSC	11	3.4	33.3	33.3
	Self employed	4	1.3	12.1	45.5
	Hospital	2	.6	6.1	51.5
	Casual	5	1.6	15.2	66.7
	Company	6	1.9	18.2	84.8
	Government	2	.6	6.1	90.9
	County government	3	.9	9.1	100.0
	Total	33	10.3	100.0	
Missing	System	286	89.7		
Total		319	100.0		

If no please indicate whether you have a source of income

		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Farming	141	44.2	53.8	53.8
	Business	75	23.5	28.6	82.4
	Farming and business	32	10.0	12.2	94.7
	Carpentry	1	.3	.4	95.0
	Masonry	1	.3	.4	95.4
	Farming and casual laborer	3	.9	1.1	96.6
	Retirement benefit	3	.9	1.1	97.7
	Mining	1	.3	.4	98.1
	Business and pension	2	.6	.8	98.9
	Business and mining	1	.3	.4	99.2
	Online writing	2	.6	.8	100.0
	Total	262	82.1	100.0	
Missing	System	57	17.9		
Total		319	100.0		

Crop Cycles in your Farm:
82% of respondents indicated two crop cycles per year, 17.3% indicated one cycle and 0.7% indicated three crop cycles.

Please indicate the number of crop cycles per year in your farm

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	47	14.7	17.3	17.3
	2	223	69.9	82.0	99.3
	3	2	.6	.7	100.0
	Total	272	85.3	100.0	
Missing	System	47	14.7		
Total		319	100.0		

What type of pest attacks do you encounter in your farming activities?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Stalk Borer	10	3.1	3.6	3.6
	Stalk borer and Weaver birds	4	1.3	1.4	5.0
	Cut worm and Stalk borer	1	.3	.4	5.4
	Stalk Borer, Aphids and Weaver birds	1	.3	.4	5.7
	Blites	2	.6	.7	6.4
	Worms	101	31.7	36.1	42.5
	Weeds	19	6.0	6.8	49.3
	Worms and weeds	28	8.8	10.0	59.3
	Rodents and worms	1	.3	.4	59.6
	Aphids	4	1.3	1.4	61.1
	Weaverbirds and aphids	3	.9	1.1	62.1
	Rodents, Weaverbirds and worms	4	1.3	1.4	63.6
	Worms and aphids	36	11.3	12.9	76.4
	Worms and weevils	13	4.1	4.6	81.1
	Stalk borer and weevils	4	1.3	1.4	82.5
	Birds, worms and monkeys	1	.3	.4	82.9
	Birds and weeds	1	.3	.4	83.2
	Monkeys and Squirrels	1	.3	.4	83.6

	Weevils	12	3.8	4.3	87.9
	Molds and worms	1	.3	.4	88.2
	Cricket, worms, birds and mice	2	.6	.7	88.9
	Birds and worms	1	.3	.4	89.3
	Locust	30	9.4	10.7	100.0
	Total	280	87.8	100.0	
Missing	System	39	12.2		
Total		319	100.0		

75% of respondents indicated that they did not use organic pest control techniques while 24% indicated that they use organic pest control techniques.

Do you use any organic pest control techniques?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	236	74.0	75.2	75.2
	Yes	78	24.5	24.8	100.0
	Total	314	98.4	100.0	
Missing	System	5	1.6		
Total		319	100.0		

If yes please specify

		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Pesticides	35	11.0	47.9	47.9
	Local herbs	3	.9	4.1	52.1
	Manure	13	4.1	17.8	69.9
	Weeding	6	1.9	8.2	78.1
	Wood ash	14	4.4	19.2	97.3
	Pesticides and herbs	1	.3	1.4	98.6
	Pesticide and ash	1	.3	1.4	100.0
	Total	73	22.9	100.0	
Missing	System	246	77.1		
Total		319	100.0		

95.5% of the respondents do not use any type of irrigation while 4.5% indicated that they use irrigation in their farming activities.

Do you use any type of irrigation?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	299	93.7	95.5	95.5
	Yes	14	4.4	4.5	100.0
	Total	313	98.1	100.0	
Missing	System	6	1.9		
Total		319	100.0		

If yes, specify

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Rainfall	16	5.0	80.0	80.0
	Drip irrigation	1	.3	5.0	85.0
	Sprinkle	3	.9	15.0	100.0
	Total	20	6.3	100.0	
Missing	System	299	93.7		
Total		319	100.0		

73.8% of the respondents indicated that they use Ox ploughing method to do farming, 2.5% use hand digging method, 9.1% use tractors and 10.7% use both oxen and tractor to plough.

What ploughing methods do you use in farming?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Ox plough	234	73.4	73.8	73.8
	Hand digging	8	2.5	2.5	76.3
	Ox plough and Hand digging	10	3.1	3.2	79.5
	Tractor plough	29	9.1	9.1	88.6
	All methods	2	.6	.6	89.3
	Oxen and tractor	34	10.7	10.7	100.0
	Total	317	99.4	100.0	
Missing	System	2	.6		
Total		319	100.0		

Appendix 6: NDVI averages

	Wkt_geom	_mean
1	Point (679440.56565792253240943 9860357.49709988757967949)	0.530378622726901
2	Point (679322.60980006842873991 9859664.1899594459682703)	0.509600371349184
3	Point (679692.43334136658813804 9856491.2468635980039835)	0.543608204119944
4	Point (680359.45041922363452613 9854840.06545592471957207)	0.538071751765194
5	Point (680988.02928306418471038 9853659.17531372234225273)	0.540530577666064
6	Point (683031.34388000785838813 9852836.48719156533479691)	0.51891357945354
7	Point (686422.71254928084090352 9852453.2730083130300045)	0.524743147588623
8	Point (689603.28217022749595344 9851039.77896921150386333)	0.553811495184186
9	Point (689916.28274229553062469 9850623.3033433835953474)	0.543480003895261
10	Point (690661.03552251192741096 9849483.75616908445954323)	0.546819193696806
15	Point (692769.64183439291082323 9847513.47706030867993832)	0.53040276924139
16	Point (683355.97616871655918658 9860876.84095096960663795)	0.502420804671969
17	Point (683837.02257441938854754 9863050.41132395341992378)	0.53192711945134
18	Point (684185.59619820269290358 9864655.58410394564270973)	0.514158010973312
19	Point (683971.26121742837131023 9865596.99114113114774227)	0.531419316457991
20	Point (680199.48004057316575199 9871730.49440791457891464)	0.555765876214926
21	Point (676244.20833044592291117 9873475.83997459150850773)	0.544046552762502
22	Point (674662.69497119239531457 9874075.65723027288913727)	0.519489399496097
25	Point (664558.39391666138544679	0.515938525515641

	9864616.71338533423841)	
26	Point (664558.39391666138544679 9864616.71338533423841)	0.515938525515641
27	Point (664558.39391666138544679 9864616.71338533423841)	0.515938525515641
28	Point (664558.39391666138544679 9864616.71338533423841)	0.515938525515641
29	Point (664558.39391666138544679 9864616.71338533423841)	0.515938525515641
30	Point (656496.39124373695813119 9867948.12466607056558132)	0.521134865580562
31	Point (659324.57909836352337152 9869850.89681601338088512)	0.541376577931797
32	Point (659480.477145736804232 9870039.19794951193034649)	0.549869677613473
33	Point (661015.16214920091442764 9872275.1567543875426054)	0.552200564150101
34	Point (660090.50818978762254119 9884592.67280906438827515)	0.51876968976664
35	Point (661555.12309530971106142 9884130.2200742345303297)	0.529750525379287
36	Point (662785.12459000339731574 9883642.56820089742541313)	0.531742553571848
37	Point (662161.55945386225357652 9879218.30068132840096951)	0.519310391387899
38	Point (660969.87315332877915353 9879715.15175362676382065)	0.564973691431332
39	Point (660554.83214728278107941 9880803.82600172236561775)	0.554575179558006
40	Point (658054.47653320198878646 9882013.95980805344879627)	0.528196832512395
41	Point (656193.93616786715574563 9883492.56594704277813435)	0.532090914939421
42	Point (653519.22756504314020276 9885053.32513583824038506)	0.515310661296357
43	Point (651141.97535762959159911 9885281.48287332430481911)	0.504946794451019
44	Point (649780.3001673761755228 9885912.68349406868219376)	0.481212133690602
45	Point (647275.130725302034989	0.457249609943555

	9886771.59273764863610268)	
46	Point (643504.63289661961607635 9887043.74422578327357769)	0.469675890694302
47	Point (642949.42829846532549709 9887321.32448149658739567)	0.428752134461533
48	Point (642736.62036091112531722 9887806.15214241109788418)	0.428752134461533
49	Point (645284.42322856199461967 9886203.99075493216514587)	0.433128469527245
50	Point (646680.92048507393337786 9883192.89295990206301212)	0.474811076795229
51	Point (648218.25622139195911586 9881712.65014345198869705)	0.453105551646306
52	Point (650787.57616785261780024 9879574.35859684646129608)	0.447976912880558
53	Point (654701.78587919357232749 9877936.23976829461753368)	0.502914923106656
54	Point (657373.62876236066222191 9879434.12171704322099686)	0.552934985162906
55	Point (666819.60728897608350962 9880164.97451196610927582)	0.562831184248856
56	Point (667001.89714943408034742 9879698.52782985381782055)	0.575207040813165
57	Point (667934.35508329828735441 9879332.93038132041692734)	0.575207040813165
58	Point (669493.367339524673298 9878901.82888627611100674)	0.574441891869582
59	Point (670637.79993730736896396 9878882.03914519958198071)	0.584988406852311
60	Point (672023.2906010610749945 9878405.88302105106413364)	0.605272136233352
61	Point (673895.05378983053378761 9879668.23590024001896381)	0.59540951910507
62	Point (675767.03452790912706405 9879809.36873546056449413)	0.545298831264978
63	Point (676582.45417499961331487 9880119.67380283027887344)	0.526762035537511
64	Point (671087.30967780144419521 9883729.51878085173666477)	0.566224548872739
65	Point (668250.56234465737361461	0.591208434607049

	9884119.035616310313344)	
66	Point (666388.98859938629902899 9884242.05218010395765305)	0.552750504011808
67	Point (656235.74729307438246906 9890019.15174872428178787)	0.522024954858538
68	Point (655985.51877126016188413 9890439.49416319280862808)	0.530130714786762
69	Point (655495.9263261912856251 9890657.18052475154399872)	0.530130714786762
70	Point (652397.53054537298157811 9893264.43793674185872078)	0.494192269654931
71	Point (654044.59067481523379683 9893126.98250356875360012)	0.499932776439098
72	Point (654351.16659109084866941 9893590.70889465138316154)	0.499932776439098
73	Point (667149.83613623818382621 9890772.75316683389246464)	0.536234371397656
74	Point (672018.39593056566081941 9890866.77629536762833595)	0.504994121429821
75	Point (672606.30331994208972901 9888789.55566420964896679)	0.542191194340738
76	Point (673977.36055191000923514 9887553.27211404033005238)	0.551389064508791
77	Point (671595.95101622748188674 9893152.13128348626196384)	0.565516276570891
78	Point (673400.35487234476022422 9894381.14506406709551811)	0.549912749456931
79	Point (673261.68327775248326361 9895637.55396575108170509)	0.477803017312051
80	Point (671363.93337106448598206 9899054.46057437360286713)	0.523672718306383
81	Point (668903.53628733905497938 9895472.48346753604710102)	0.563534362031124
82	Point (667768.73737877840176225 9895228.97451860457658768)	0.56284224219645
83	Point (666522.49971491657197475 9895934.71968108788132668)	0.544256967437111
84	Point (664298.85689878044649959 9899621.38853724487125874)	0.541393243350632
85	Point (663104.46793678076937795	0.534247556977448

	9899772.97950156778097153)	
86	Point (664308.84798820503056049 9900104.49163486622273922)	0.54244514138383
87	Point (668502.45713091862853616 9900542.31657341681420803)	0.457297504236731
88	Point (668473.22848261729814112 9900646.90503386408090591)	0.457297504236731
89	Point (669528.63824441051110625 9900248.73050721548497677)	0.519377014919883
90	Point (671341.56030625582206994 9901951.04616396315395832)	0.57325878756303
91	Point (671368.43043318146374077 9902330.55170716717839241)	0.54559257554695
92	Point (671690.26210998254828155 9902735.49162381701171398)	0.54559257554695
93	Point (674230.12988515058532357 9905252.39481383003294468)	0.527663307883768
94	Point (673948.39084334822837263 9906005.74656268954277039)	0.539504153068832
95	Point (673915.89684586285147816 9907129.38932887651026249)	0.540558197039073
96	Point (669463.65484678791835904 9897403.5407955851405859)	0.40835378070029
97	Point (669332.985842494526878 9896623.91550289653241634)	0.326982808966883
98	Point (669282.04071675823070109 9896436.85385649465024471)	0.326982808966883
99	Point (677052.75836148613598198 9902447.38884106278419495)	0.518209151360256
100	Point (677459.25690694828517735 9901699.52167687937617302)	0.527819915864815
101	Point (678933.92705252906307578 9900143.19211105071008205)	0.559208910241672
102	Point (676445.48308113496750593 9908383.18475032597780228)	0.53333774428327
103	Point (677013.74679793370887637 9909610.33684686385095119)	0.563332594518527
104	Point (677806.1593193932203576 9913076.55710132978856564)	0.561807316837995
105	Point (680837.93375815311446786	0.58028982958638

	9914660.8668854720890522)	
106	Point (680803.44536988995969296 9914416.23839645087718964)	0.58028982958638
107	Point (680971.84303720598109066 9916137.60010997392237186)	0.607414475335324
108	Point (681242.41689733287785202 9919395.89534517005085945)	0.598332263123576
109	Point (681132.41009891766589135 9919926.93465329706668854)	0.598332263123576
110	Point (681192.61394038470461965 9920193.76850984059274197)	0.561128241963246
111	Point (673082.15935891144908965 9922161.27258048951625824)	0.508889183129781
112	Point (673946.08558531454764307 9921689.48848781734704971)	0.522423452025024
113	Point (674548.03409251268021762 9921340.37537724897265434)	0.560256238182899